

```

Db 1042 ACCTGCTCTTGGCCGCTGCTGTTATATAAAGATCTTACCTGCTTACGGGCGCTC 1101
Qy 361 ALAAspGlyLeuValAlaAlaPheProValValArgGlyThrProLysAspSerCysSer 380
Db 1102 GCCGATGGGGCTGTGGCTGTGTTCCCTGGTGGCGGGCACCCCAAGAGACAGCTGCTCC 1161
Qy 381 TyrlaucysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
Db 1162 TACCTGTCTCTACACACAGCCACAGCTCCAGTTCCAGCTCCGGATGAGAACCCAGCG 1221
Qy 401 GluAsnProTyrProValLysAlaMetGluValValAsnSerGlySerGluValTyrPyr 420
Db 1222 CAGAACCCCTACCCAGTGAAGGAGGAGGTGTCACAGAGGGCTCGAGGTCTGTGTAC 1281
Qy 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
Db 1282 AGCAATGGGGCGGGCTCTGTCTCATCGACTGTGCTCCCTGGAGATCTGCAGGGCGCTG 1341
Qy 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGluValArgGly 460
Db 1342 GAGCCCTACATGCCCCCTCCATGTTACGTACGTGCTGTGAGCTCTGAGGGCAGAGGG 1401
Qy 461 GluGluValValTyrCysLeuAspAspLysAlaAsnSerLeuValMetLysHisSerThr 480
Db 1402 GAGGAGGTGCTGTGCTGTGATGACAGGCAACCTCCTGTGTATGACACCTCACACC 1461
Qy 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
Db 1462 ACCTACCAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1521
Qy 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
Db 1522 CCCGTGCGGCTTGTGACACGAGAACCCCGGACGACCCACAGCCCAACCAAGGTG 1581
Qy 521 ProGluGlyAspSerIleAlaAspValSerIleMetLysSerGluGluLeuGlyThrGln 540
Db 1582 CCTGAGGGGAGACTCATGCGGAGCTGAGCAGCATGATGAGAGAGAGTGGGACAGCG 1641
Qy 541 IleLeuIleHisGlnGluSerLeuThrAspTyrCysSerMetSerSerLysSerSerSer 560
Db 1642 ATCCGATTCACACAGAGATCATCTACTGCTCTCATCTCTCTCTCTCTCTCTCTCTCT 1701
Qy 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
Db 1702 CCACCCCGCAGGCTGCCAGGTGCCCTCTCAAGCTCCCGAGTCCCGCAGCAAGTCTTCC 1761
Qy 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
Db 1762 AGTGTGCTTCTCTCCACGACGACTGCGAGACTCAGACATGCTACATACGCCCGGTCTGC 1821
Qy 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
Db 1822 TCCGACAGGCTGTGACATGACTGACCCCATGAGCGGAGGAGACCTTCAGCAGCACCTGG 1881
Qy 621 GlnAlaValLysIleLeuAlaValArgAspLeuIleTyrValProArgArgGlyGlyAsp 640
Db 1882 CAGGCGGTGAATATCTCGCCCTCAGAGACTCATTTGGGTGCCAGGGCGGTGGAGAT 1941
Qy 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyValIleAlaVal 660
Db 1942 GTTATGCTGATGGCTGTGAGAGATCTGTGAAGCCAGCGGGCGGAGTCTATTGCGCTG 2001
Qy 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValValAlaLys 680
Db 2002 TTAAGAAGCCGAGAGCTGCTCCGATGGGTGGGTGTGTGTGTGTGTGTGTGTGTGTGT 2061
Qy 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrCysLeuAlaValTyrArg 700
Db 2062 GACACTGTGTGTGACCTTTGAATAAGAAACAGAGATGTGTGTGTGTGTGTGTGTGTG 2121
Qy 701 GlyTyrGlyAlaArgGluPheAspIlePheTyrGlnSerTyrGluGluLeuArgLeu 720

```

```

Db 2122 GCGTGGGCGCCAGGAGTGTGACATTTCTACAGTCTACAGAGAGCTGGCGGCTG 2181
Qy 721 GlnAlaCysThrArgLysArgArg 728
Db 2182 GAGGCTTGACCTCGCAAGAGAGG 2205

RESULT 2
US-09-836-392-8
Sequence 8, Application US/09836392
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
FILE REFERENCE: PTO20P1
CURRENT APPLICATION NUMBER: US/09/836,392
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: PCT/US00/28066
PRIOR FILING DATE: 2000-10-11
PRIOR APPLICATION NUMBER: 60/159,542
PRIOR FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: 60/165,914
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: 60/189,027
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 3496
TYPE: DNA
ORGANISM: Homo sapiens
US-09-836-392-8

Alignment Scores:
Pred. No.: 5.56e-267 Length: 3496
Score: 3822.00 Matches: 728
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-836-392-21 (1-728) x US-09-836-392-8 (1-3496)
Qy 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 20
Db 22 ATGCTGAGGACCACTCGGGGCGCCAGCATGACAGAAATCTTCGAGTCCGGGAGAGG 81
Qy 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
Db 82 GCCACATGCTGACAGCGGCTGCAGCACCCCTCATCGTGGCCTCATCGGCATCGACATC 141
Qy 41 HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
Db 142 CACCCGCTGTGCTTGGCCCTGAGCTCGCGCTCAGAGCTCAACACCTGCTGTCTC 201
Qy 61 GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnLysIleAla 80
Db 202 GAGAACGCGACAGATTTCTCTTATACCCCTGGGACACATGCTCACCCAAAAATAGCC 261
Qy 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
Db 262 TACCAAGATCGCTCGGGCTGTGCTTACCTGACCAAGAAAAACATCATCTTCTGTGACCTG 321
Qy 101 LysSerAspAsnIleLeuValTyrSerLeuAspValLysGluHisIleAsnIleLysLeu 120
Db 322 AGTGTGACACATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 381
Qy 121 SerAspTyrGlyIleSerArgLysPheHisGluGlyAlaLeuGluValGluGlyThr 140
Db 382 TCTGACTACGGGATTTGAGGACATCATCATGAGGGGCGCCCTAGCGCTGAGGGCACT 441
Qy 141 ProGlyTyrGlnAlaProGluLysArgProArgIleValItyrAspGluLysValAspMet 160
Db 442 CCTGCTACCAAGCCCGCAGATCAGGCTCGCATTTGTATGATGAGAGAGTAAATAG 501

```

QY 161 PheserTyrGlyMetValLeuTyrGluLeuSerGlyGlnArgProAlaLeuGlyHis 180
 |||||
 Db 502 TTCTCCATGGAATGGTCTCTACAGATGCTGTCCAGAGACGCCCTGCACTGGGCGAC 561
 QY 181 HisGluLeuGlnIleAlaValLeuSerGlyGlyLeuArgProValLeuGlyGlnPro 200
 |||||
 Db 562 CACCACTCCAGATGCGCAAGAGCTGTCCAAAGGGCATCCGCGGTTCTGGGGCAGCCG 621
 QY 201 GluGluValGlnPheArgArgLeuGlnAlaLeuMetGlyCysTyrPhePhePro 220
 |||||
 Db 622 GAGAAATGCAATGCGCGGCACTGCAAGCGCTCATGATGAGTCTGGGACACTAAGCA 681
 QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
 |||||
 Db 682 GAGAAAGGACCGCTGCGCTGCTGAGAGCAGATGAAGAGACCCGACTTTGGCCACC 741
 QY 241 PheMetTyrGluLeuGlyCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGln 260
 |||||
 Db 742 TTCATGTATGAACGTCTGTGGGAGACAGACACCTTCTCTATCCAGGGCCAGAG 801
 QY 261 TyrThrValValPheTyrPheArgLysGluGlnSerArgAsnTyrThrValValAsnThr 280
 |||||
 Db 802 TACACCTGGTGTGGTGGATGAGAAAGAGAGAGTCCAGAACTACAGGTTGTGAACCA 861
 QY 281 GluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
 |||||
 Db 862 GAGAAAGGCGCTCATGAGAGTGCAGAGATGTCTCCCTGGGATGAAGTGAAGTGCAG 921
 QY 301 LeuGlnValGlnArgSerLeuTyrThrAlaThrGlnAspGlnLysIleTyrThr 320
 |||||
 Db 922 CTCAGAGTCCAGAGATCCCTGTGGACAGCCAGAGAGCCGAAATCTCATATCTACACC 981
 QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaVal 340
 |||||
 Db 982 CTCAAGGCGAGTGCCTTAAACACACCCCAAGCGCTTGATCTCCAGCTGTCTC 1041
 QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
 |||||
 Db 1042 ACCTGCTTCTGGCGCGCTGTATTAAGAAATCTCTACGTGTGAGCGGCGCTC 1101
 QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 |||||
 Db 1102 GCCCATGGGCTGTGGTGTGTCTCCGTGTGGGGGACCCCAAGAGACGCTGCC 1161
 QY 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlnAspAlaArg 400
 |||||
 Db 1162 TACCTGTGCTACACACAGCCACAGGTCAGATTCAGCATCGCGATGAAGACGACAG 1221
 QY 401 GlnAsnProTyrProValLysAlaMetGluValValAsnSerGlySerGluValTyrPyr 420
 |||||
 Db 1222 CAGAAACCCCTACAGAGAGAGGAGGAGTGTGTCAGACAGCGGTCTGAGTGTGAGTAC 1281
 QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
 |||||
 Db 1282 AGCAATGGGCGGCGCTCTGTGTATCATGACTGTCTCCCTGGAGATCTGAGGGGCGTG 1341
 QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgGly 460
 |||||
 Db 1342 GAGCCCTACATGCGCCCTCATGTTACGTACGTGTGTGAGTCTGAGGGGCGAGAGG 1401
 QY 461 GluGluValValTyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 |||||
 Db 1402 GAGGAGGTCGT 1461
 QY 481 ThrTyrGlnLeuGlyAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
 |||||
 Db 1462 ACCACACAGCTGT 1521
 QY 501 ProValArgProLeuAspThrGluProProAlaIleSerHisThrAlaAspProLysVal 520
 |||||
 Db 1522 CCGGTGGGCGCTGT 1581

QY 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluGluGlyThrGln 540
 |||||
 Db 1582 CTTGAGGGGAGCTCCACCTGCGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1641
 QY 541 IleLeuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 560
 |||||
 Db 1642 ATCTGTATCCACAGAAATATCTACTGACTACTGCTCATGTCTCTCTCTCTCTCTCTCT 1701
 QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
 |||||
 Db 1702 CACACCCCGAGGCTGCGAGGCTCCCTCAAGCCTCCACAGCTCCACAGCAAGTCTTCC 1761
 QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
 |||||
 Db 1762 AGTGTCTCTTCTCCACCGAGCTGCGAGAGCTCAACATGCTATACAGCCGCTGCTGCC 1821
 QY 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
 |||||
 Db 1822 TCCAGAGGCTGTGAGTATGACTGATCCCTTGGAGCGGGAGACCTTCAGCCAGCAGCTG 1881
 QY 621 GlnAlaValLysIleLeuAlaValArgAspLeuIleTyrValProArgArgGlyLysPhe 640
 |||||
 Db 1882 CAGCGCGTGAAGATCTCGCGCTGAGAGACCTATTTGGTCCCGCAGCGCGGTGAGAT 1941
 QY 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 |||||
 Db 1942 GTTATCTCATTTGGCTGTGGAGAGATTCGAAACCCAGCGGGCCAGATCATTTGGCTC 2001
 QY 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 680
 |||||
 Db 2002 TTTAAAGCCGAGAGCTGACTCCCATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2061
 QY 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrCysLeuAlaValTyrPhe 700
 |||||
 Db 2062 GACACTGTGTGTGCTACCTTGAATAATGAATAACAGAGAGTGTCTGCGCGCTGTGAGG 2121
 QY 701 GlyTyrPheLysAlaArgLysPheAspIlePheTyrGlnSerTyrGluGluLeuArgLeu 720
 |||||
 Db 2122 GCGTGGGCGCGCAGGAGGATTCGACATTTTCACAGTCTTACAGAGAGTGGCGGCTG 2181
 QY 721 GluLacCysThrArgLysArgArg 728
 |||||
 Db 2182 GAGGCTTGCACTCCAAAGAGAGAG 2205
 |||||
 RESULT 3
 PCT-US01-01435-37
 : Sequence 37, Application PC/TUS0101435
 : GENERAL INFORMATION:
 : APPLICANT: Human Genome Sciences, Inc.
 : TITLE OF INVENTION: 21 human secreted proteins
 : FILE REFERENCE: PS726PCT
 : CURRENT APPLICATION NUMBER: PCT/US01/01435
 : PRIOR FILING DATE: 2001-01-17
 : PRIOR APPLICATION NUMBER: 60/226,282
 : PRIOR FILING DATE: 2000-08-18
 : NUMBER OF SEQ ID NOS: 145
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 37
 : LENGTH: 3579
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : PCT-US01-01435-37
 Alignment Scores:
 Pred. NO.: 5,71e-267 Length: 3579
 Score: 3822.00 Matches: 728
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0
 US-09-836-392-21 (1-728) x PCT-US01-01435-37 (1-3579)

QY 1 MetLeuAlaGHisLeuAlaAlaThrAspAlaMetLysAsnPheserGluPhargInGlu 20
DB 68 AGCGTAGGCGACCTGGCGGCGACCGATGCCATGGAAGAACTTCGAGATTCGCGAGAG 127
QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
DB 128 GCCAGACATGCTCAGCGGCTGCAACACCCCTGATCGTGCGCTATCGGATCAGACATC 187
QY 41 HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
DB 188 CACCGGCTGCTGCGCTGAGCGCTGCGCGCTCAGACGCTCAACACCGCTGCTGCC 247
QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGluHisMetLeuThrIleAlaIle 80
DB 248 GAGAACGCCAGAGATTCCTCTTATACCTCGGAGACATGCTCAACCAAAAATAGCC 307
QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
DB 308 TACCAAGATCGCTCGGCGCTGCGCTACCGCAACAGAAAAATCATCTTCGTGACCTG 367
QY 101 LysSerAspAsnIleLeuValIlePheSerLeuAspValLysGluHisIleAsnIleLysLeu 120
DB 368 AAGTGGACAAACATCTGGTGTGTGCTTGTACGTCAGAGACACATCAACATCAAGCTA 427
QY 121 SerAspTyrGlyIleSerArgLysSerPheHisGluGlyAlaLeuGlyValGluGlyThr 140
DB 428 TGTGACTAGCGGATTCGAGGCGAGTCATTCATGAGGGGCGCTTGGGCGGCGACAT 487
QY 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet 160
DB 488 CTTGGCTACACGAGCCCGCAAGATCAGGCTCCCATTTGTATGATGAGAAAGTAGATAG 547
QY 161 PheSerTyrGlyMetValLeuTyrGluLeuLeuSerGlyLysArgProAlaLeuGlnHis 180
DB 548 TTCTCTATAGATAGGTGCTTACGAGTGTGCTGAGAGACAGCGCTCGACAGGGCCAC 607
QY 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyLysPro 200
DB 608 CACCAAGCTCAGATTCGCAAGAGCGTCGCAAGGGCGATCGCGCTTGTGGGCGCGG 667
QY 201 GluGluValAlaGlnPheArgTyrLeuGlnAlaLeuMetMetGluCysArgPheAspThrLysPro 220
DB 668 GAGGAAGTCACTTCGCGGAGCTGCGGCGCTCATGATGATGCTGCGGACACTAAGCCA 727
QY 221 GlyLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
DB 728 GAGAAAGCGACCGCTGCGCTGCGGTGAGAGACAGATGAAAGACCGACTTTTGCACCC 787
QY 241 PheMetTyrGluLeuCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGlu 260
DB 788 TTCATGATATGAACCTGCTGTGGAGAGACAGACGCTTCTTCATCCACAGGCGCGAG 847
QY 261 TyrThrValValPheThrAspGlyLysGluLysArgAsnTyrThrValValAsnThr 280
DB 848 TACACCGTGTGTTTGGGATGGAAGAGAGATGCTGCTGCGATGAAAGTAGAGCTCCAG 907
QY 281 GlyLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
DB 908 GAGAAAGGCGCTCATGAGAGGAGAGAGATGCTGCTGCGATGAAAGTAGAGCTCCAG 967
QY 301 LeuGlnValAlaGlnArgSerLeuThrPheAlaThrGluAspGlnLysIleTyrIleThr 320
DB 968 CTCACAGTCCAGAGATCCCTGTGGACAGCCAGCAGAGACAGAAATCTCACTCATCACC 1027
QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAla 340
DB 1028 CTCAGAGGATGTGCGCTTAAACACACCCAGAGCGCTTGATATCTCAGTGTGCTGCTC 1087
QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
DB 1088 ACCTGCTTCTGGCGCTGCTGTATTAAGAAATTCCTACCTGCTTACGAGGCGCTC 1147
QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380

DB 1148 GCCGATGGGCTGTGTGCTGTGTTCCCTGTGCGGCGACCCCAAGACAGCTGCTCC 1207
QY 381 TyrLeuCysSerHisIleThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
DB 1208 TACCTGTGCTGCACACAGAGCCACAGTCCAGATGCGGAGATGAAAGACGAGG 1267
QY 401 GluAsnProTyrProValLysAlaMetGluValValAsnSerGlySerGluValIlePhe 420
DB 1268 CAGAACCCCTTACCGAGTGAAGGCCATGAGGTGCTCAACAGCGGCTGAGCTGTGATAC 1327
QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
DB 1328 ACCAATGGCGGCGCTGCTGTGCTGATGAGCTGCTGCTGAGATCTGCAAGGCGCTG 1387
QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGluLysArgGly 460
DB 1388 GAGCCCTCATGCGCCCGCTCCATGTTAGCTAGCTGTGCAAGCTGTGAGGCGAGAGG 1447
QY 461 GluGluValValTyrCysLeuAspLysAlaAsnSerLeuValMetLysHisSerThr 480
DB 1448 GAGGAGTGTGCTGCTGCTGCTGATGACAGGCAACTCTTGTGTATGATCCTCACAC 1507
QY 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
DB 1508 ACCTACCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1567
QY 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisIleThrAlaAsnProLysVal 520
DB 1568 CCGGTGCGGCTTGTGACAGCGAAGACCCCGGAGCGACAGCCAGCCAAAGGTTG 1627
QY 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluGluLeuGlyThrGln 540
DB 1628 CCGTGGGCGGACTCCATCGCGAGCTGAGCATCATGATACAGAGAGAGCTGGGCGACG 1687
QY 541 IleLeuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerSerSerSerSer 560
DB 1688 ATCTGATTCACAGAGAACTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1747
QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
DB 1748 CCACCGCGCAGGCGCGAGTCCCGCTCAAGCTCCCGACGCTCCCGACAGTTCTTCC 1807
QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
DB 1808 AGTGTGCTTCTTCCACACCACTGCGAGGACTCAGACATGCTCATACGCGCTGCTGCC 1867
QY 601 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
DB 1868 TCCGACAGGTCTGAGCATGACCTGACCCCATGAGAGGAGAGACTTCAGCCAGCACTG 1927
QY 621 GlnAlaValLysIleLeuAlaValArgAspLeuIleTyrValProArgArgGlyLysAsp 640
DB 1928 CAGCGCGTAAATCTCCCGCTCAGAGACATCATTTGGGTCCCGAGCGGCGGTGAGAT 1987
QY 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
DB 1988 GTTATGCTGATTTGGCTGAGAGAGATCTGAGAGCCCAAGCGGCGGAGTCTATTCGCTC 2047
QY 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 680
DB 2048 TTAAGAGCCCGAGAGGTGACTCGGCTGAGGAGGCTGAGTCCCTGTGGGAGAAAG 2107
QY 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrCysLeuAlaValTyrArg 700
DB 2108 GACACGTGTGTGCTGCTGCTTGAATAATGAAGACAGAGTGTGTGCTGCGCTGTGAGG 2167
QY 701 GlyTyrPylAlaArgLysPheAspIlePheTyrGlnSerTyrGluLysLeuGlyArgLeu 720
DB 2168 GCGTGGGCGCGCAGGAGATTCGACATTTCTACACATGCTACAGAGAGAGCTGGCGCGCTG 2227
QY 721 GluAlaCysThrArgLysArgArg 728

Db 2228 GAGGCTTGCACTCGCAGAGAAGG 2251

RESULT 4

PCT-US01-01435-14
: Sequence 14, Application PC/TUS0101435

; sequence 14, Application PC/1050101433
; GENERAL INFORMATION:

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: 31 human secreted

FILE REFERENCE: PS726PCT

CURRENT APPLICATION NUMBER

PRIOR APPLICATION NUMBER: 60/226,382

PRIOR FILING DATE: 2000-08-18

NUMBER OF SEQ ID NOS: 145

; SOFTWARE: PARENTHL.VEL. 2
; SEQ ID NO 14

LENGTH: 3564

ORGANISM: Homo sapiens

PCT-US01-01435-14

21425000 + 6000000

Alignment scores:
Pred. No.: 1.

Score: _____

Alignment Scores:

Pred. No.:	1,83e-266	Length:	356
Score:	.3815	Matches:	727
Percent Similarity:	99.86	Conservative:	0
Best Local Similarity:	99.86	Mismatches:	1
Query Match:	99.828	Indels:	0
DB:	1	Gaps:	0

US-09-836-392-21 (1-728) x PCT-US01-01435-14 (1-3564)

Oy	1	MeeuuarGHSleuAaGalaThrAspAlaMetLysanbHeserGluPhearGyGlnGlu	20
Dd	74	ATGCTGAGGACACTCGGGGGCCACCGAGATGCCATGAAGAATCTTCGCGATTCCGGCAGAG	133
Oy	21	AlaserMetLeuHsAlaLeuGlnHsProCysIleValAlaLeuIleGlyIleSerIle	40
Dd	134	GCCAGCAATGCTGCACGGGGCTGCACACCCCTGCATCTGGCGGCATCGGCATCAGCATC	193
Oy	41	HsProLeuGlyPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer	60
Dd	194	CACCCGCTCTCTTCGGCCCTGGACCTGGCCGCTTCAGCAACCTCAACCGTGGTCTCC	253
Oy	61	GluAsnAlaArGAsPserSerPheIleProLeuGlyHsMetLeuThrGlnIleAla	80
Dd	254	GAGAAAGCCAGAGATCTCTCTTTATACCCCTGGGACACATGCTCACCCAAAAATAGCC	313
Oy	81	TyrGlnIleAlaSerGlyLeuAlaIleTyrLeuHsLysLysAsnIleIlePheCysAspLeu	100
Dd	314	TACCAAGATCGCCTGGCGGCTGGCTACCTGCACAAAGAAAAACATCATCTTCTGTGACCTG	373
Oy	101	LysSerAspAsnIleLeuValTrpSerLeuAspValLysGlnHsIleAsnIleLysLeu	120
Dd	374	AAGTGGACACATCTGGTGGTGGTCCCTTAGCTCAAGGAGGACATCAACATCAAGCTA	433
Oy	121	SerAspTyrGlyIleSerArgGlnSerPheHsGluGlyAlaLeuGlyValGluGlyThr	140
Dd	434	TCTGACTAGGGGATTTCGAGGCACTCATCCATAGGGCGCCCTCAAGCGGTGGAGGAGACT	493
Oy	141	ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet	160
Dd	494	CCTGGCTACAGGGCCCCAGACATCAGGGCTGGCATGTATATGATGAAAGTGAATG	553
Oy	161	PheSerTyrGlyMetValLeuTyrGlyLeuLeuSerGlyGlnArgProAlaLeuGlyHs	180
Dd	554	TTCTCTCATGGAAGGTGCTCTACGAGTGTGTGCAGAGACGGCCCTGGACACTGGGGCAC	613
Oy	181	HsGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnPro	200
Dd	614	CACGAGCTCCAGATTGCCAABACCTGTCCAAAGGCACTCGCCCGGTTCGTGGGGCACCG	673
Oy	201	GluGluValGlnPheArgArgLeuGlnAlaLeuMetCysGlySerPheThrLysPro	220

D	674	GAGAAATGCAAGTCCGGGAGACTGACAGGGCGCTCATGATGAGAGTGGCTGGAGCACTAAACCA	733
Q	221	GLULYSARGPROLEUALALEUSERVALVALSERGINMETLYSASPPTOTHPHAIATTHR	240
D	734	GAGAAAGGACCGGTGGCCCTGTGTGGTGGAGCCAGATGAAGAGCAACCCACTTTGGCCACC	793
Q	241	PHMETRYICLULEUCYSCYSGLYLVSGLINTRRALAPHEPSESERGLIGLYLGLU	260
D	794	TTCTAGTATACACTGTGCTGTGGGAAGCAGAACGCCCTTCTCTATCCCAAGGCCAGAG	853
Q	261	TYTTHRVALVALPHEPTRAPSPGLYLSGLUGLUSERARGASNTYTRHVALVALASNTHR	280
D	854	TACACCGTGGTGTGGTGGATGGAAAGAGAGTCCAGAACTACAGGTGTGTACACA	913
Q	281	GLULYSGLYLEUWETGLUVALGINARGMETCYSCYSPROGLYMETLYSVALSERCYGLIN	300
D	914	GAGAAAGGCGCTCATGTGAGGTGCAGAGAGATGTGCTGCCGTGGANTMAAGTACGCTGCAG	973
Q	301	LEUENVALBALINPARGSERLEUTPTRRALATHRGVALASPGLINLYLLETYRILETPTHR	320
D	974	CTCCAGGTCCAGAGATCCCTGTGGAGACGACCAGAGACAGAAATATCTACATCTACACC	1033
Q	321	LEULYSGLYMETCYSPROLEUASNTHRPROGLINNALALEUASPTHRPROALAVALVAL	340
D	1034	CTCAAGGGCACTGTGCCCCCTTAAACACACCACCAAGCGCTTGGATACCTCAGCTGTGTC	1093
Q	341	THRCYPHELEUALAVALPROVALILELYLSYASNSERTYRLEUVALLEUALGLYLEU	360
D	1094	ACCTGCTTCTWTGGCGCGCTGTATTAAAGAAATCTCTACCTGTGTGTAGCGGCGCTC	1153
Q	361	ALASPGLYLEUVALALAVALPHEPROVALVALARGLYLTHPROLYSASPSECYSER	380
D	1154	GCCCATGGCGCTGTGGCTGTGTTCCTCCGTGTGGCGGGGACCCCAAGACAGCGCTGCC	1213
Q	381	TYRLEUCYSSERTHTRTALAASARGSERLYSPHESERLLEALASPGLIASPVALARG	400
D	1214	TACCTGTGCTCACACACACGCCAACAGTCCAGTTCCAGATCGCGGATGAAGACGACAGG	1273
Q	401	GLIASNPROTYRPROVALLYSALAMETGLUVALVALASNSERGLYSERGLUVALITRPTYR	420
D	1274	CAGAAACCCCTAACCCAGTGAAGGCCATGAGAGTGTCAACAGCGGCTCTGAGGTGTGGTAC	1333
Q	421	SERASNGLYPROGLYLEUVALVALLIENAPCYSAIASERLEUGLULILECYARGARGLEU	440
D	1334	AGCATGTGGCGCGGCCCTCTTGCATCGACTGTGCTCCCTGGAATCTGCAGGCGGCTG	1393
Q	441	GLUPROYMETALAPROSERMETVALITRSEVALVALCYSSERSEGLUGLYARGGLY	460
D	1394	GAGCCCTACATAGGCCCTCCATAGTTAGCTAGTGTGTGCAGCTCTGAGGGCAGAGGG	1453
Q	461	GLUGLUALVALITRACYLEUASAPPLYLSALASNSERLVALMETYRHLISERTHR	480
D	1454	GAGGAGGTCGTGTGTGTCCTGATGACAAAGGCCAACTCTTGATGTATGACACTCCACC	1513
Q	481	THRTYRGLINEUCYSAIAARGTYRPHIECYSGLYVALPROSETRPROLEUARGSPMETPHE	500
D	1514	ACCTACACAGCTGTGTGCGCGGTACTTCTGTGGGGTGTCCACAGCCCTCAGAGCACTGTTT	1573
Q	501	PROVALARGPROLEUASPTRRGUPROPROALIASERHISTHVALASNPROLYSVAL	520
D	1574	CCCGGGGGCGCTTGGACAGGAACCCCGGACGCCACGACACAGCCCAACCCAAAGGTG	1633
Q	521	PROGLUGLYASPERTLEALASPVALSERILEMETYRSERGULILEUGLYLTHGLIN	540
D	1634	CCTGAGGGGAGCTCATCTGGGAGAGTATGATATGACAGTGAAGAGTGGGACGAG	1693
Q	541	ILELEULIENHSGINGLUSERLEUTHRAPTYRCYSSERMETSESERTYRSESESER	560
D	1694	ATCTGATATCCACAGGATACACTACATGACTACTCTCATCTCTCTACTACTCCATCC	1753
Q	561	PROPROARGINALIAAARGSERPROSETRSELEUPROSETRPROALASERSETR	580
D	1754	CCACCCCGCCAGGTGTGGCAGGTCCCCCTCAAGCTCCACGATCCCAAGCAAGTTCTTCC	1813

QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGluAlaAla 600
 DB 1814 AGGTGCTCTTCTCCACCCAGCTGCGAGACTCAGACATGCTACATACGCCGGTGTGCTGCC 1873
 QY 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
 DB 1874 TCGGACAGGTCTGAGCATACCTGACCCCGACGAGGAGGAGGAGCTTCAGCCGACGACCTG 1933
 QY 621 GlnAlaValIleLeuAlaValAlaArgAspLeuIleThrValProArgArgGlyGlyAsp 640
 DB 1934 CAGGCGGTGAGATCTCTCCCGCTCAGACAGACCTCATTTGGGTGCCCGGCGGTGAGAT 1993
 QY 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 DB 1994 GTATATGCTATTTGGCTGAGAGAGATCTGGCGCCAGCGGAGCGAGTCAATGGCCGTC 2053
 QY 661 LeuLysAlaArgGluLeuThrProHisGlyValIleValAlaAlaValAlaValLys 680
 DB 2054 TTTAAAGCCCGAGAGCTGACTCCGCTGAGGGGTGCTGTGATGCTGCGCTGTGGTGAAG 2113
 QY 681 AspThrValValCysThrPheGluAsnGluAsnThrGluThrPcysLeuAlaValIlePArg 700
 DB 2114 GACAGCTGTGTGACCTTTGAATGAAACACAGAGTGTGTGCTGCGCTGCGCTGTGAGG 2173
 QY 701 GlyThrGlyAlaArgGluPheAspIlePheTyrglnSerTyrglnGluLeuGlyArgLeu 720
 DB 2174 GCTGTGGGCGCCAGGAGTTCACATTTCTACAGATCTTACAGAGAGTGGGCGCGCTG 2233
 QY 721 GluAlaCysThrArgLysArgArg 728
 DB 2234 GAGGCTGCATCGCAGAGAGAG 2257
 RESULT 5
 PCT-US01-08631-8042
 Sequence 8042, Application PC/TUS0108631
 GENERAL INFORMATION:
 APPLICANT: Hyseq, Inc
 TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 FILE REFERENCE: 21272-049
 CURRENT APPLICATION NUMBER: PCT/US01-08631
 CURRENT FILING DATE: 2001-03-30
 PRIOR APPLICATION NUMBER: 09/540,217
 PRIOR FILING DATE: 2000-03-31
 PRIOR APPLICATION NUMBER: 09/649,167
 PRIOR FILING DATE: 2000-08-23
 NUMBER OF SEQ ID NOS: 60736
 SOFTWARE: Custom
 SEQ ID NO 8042
 LENGTH: 2652
 TYPE: DNA
 ORGANISM: Homo. sapiens
 FEATURE:
 NAME/KEY: SIMILAR
 LOCATION: (136)..(1515)
 OTHER INFORMATION: 298 homologous to Caenorhabditis elegans contains similarity
 OTHER INFORMATION: to protein kinases (Pfam:PKinase_hmm, score: 149.36), accession
 OTHER INFORMATION: number AF098504, Smith-Waterman Score=427.
 PCT-US01-08631-8042
 Alignment Scores:
 Pred. No.: 2,696-260 Length: 2652
 Score: 3728.00 Matches: 719
 Percent Similarity: 98.49% Conservative: 8
 Best Local Similarity: 98.49% Mismatches: 0
 Query Match: 97.54% Indels: 3
 DB: 1 Gaps: 0
 US-09-836-392-21 (1-728) x PCT-US01-08631-8042 (1-2652)
 QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 20
 DB 445 AGCTGAGGACACCTGCGGCGCCACCGATGCGCATGAGAACTTCTCCAGTTCGCGGAGAG 504

QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
 DB 505 GCGAGATGCTGACGCGCTGACGACGCCCTGATGTGGCTCATTCGCAATCAGACATC 564
 QY 41 HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 DB 565 CACCCCTGTGCTTCCCTGAGCTCGGCGCTCAGACGCTCAACACCGGTGTGCTCC 624
 QY 61 GluAsnAlaArgAspSerPheIleProLeuGlyHisMetLeuThrGlnIleAla 80
 DB 625 GAGAACGCCAGATTTCTTCTTATACCTTGGGACATGCTCAACCAAAATATAGCC 684
 QY 81 TyrglnIleAlaSerGlyLeuAlaTyrglnHisLysAsnIleIlePheCysAspLeu 100
 DB 685 TACAGATGCGCTCGGCGCTGGCTGACATGACAAAGAAACATATCTTGTGTGACCTG 744
 QY 101 LysSerAspAsnIleLeuValIlePheSerLeuAspValLysGluHisIleAsnIleLysLeu 120
 DB 745 AAGTCGACACATCTGTGTGTGCTGCTGACGTCACAGAGACATCAACATCAAGCTA 804
 QY 121 SerAspTyrglyIleSerArgInsPheHisGlyValAlaLeuGlyValGluGlyThr 140
 DB 805 TCTGACTACGGATTTTCAGGACATGATTCATGAGAGGCGCTAGCGCTGAGGAGCT 864
 QY 141 ProGlyTyrglnAlaProGluIleArgProArgIleValTyrglnAspGluLysValAspMet 160
 DB 865 CCGTGTACACAGGCCCGACAGATCAGAGCTCGCATTTGTATGATAGAGAGATAG 924
 QY 161 PheSerTyrglyMetValLeuTyrglnLeuLeuSerGlyGlnArgProAlaLeuGlyHis 180
 DB 925 TTCTCTATGAGATGTGTCTTACGAGTGTGCTGACGAGAGGCGCTGACCTGAGGCGCAC 984
 QY 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProAlaLeuGlyGlnPro 200
 DB 985 CACCAAGCTCAGATTCAGAGAGGCTGTCAGAGGAGATCGCCGCTTGTGGGCGCGG 1044
 QY 201 GluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGluCysTrpAspThrLysPro 220
 DB 1045 GAGGAGTGCATTCGCGGAGCTGACGAGCGCTCATATGAGAGTCTGGGAGACTAAGCCA 1104
 QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
 DB 1105 GAGAAAGCGACCGCTGCGCTGCGGAGGAGACATGAGAGACCGCACTTTTCCACAC 1164
 QY 241 PheMetTyrglnLeuLysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGlu 260
 DB 1165 TTCAATGTATGAATGCTGTGTGGAGAGACAGACCTTCTTCATCCAGGCGCCAGGAG 1224
 QY 261 TyrThrValValPheThrAspGlyLysGluGlnIleArgAsnTyrglnValAlaAsnThr 280
 DB 1225 TACACCGTGTGTGTGTGGATGAGAAAGAGAGTCCAGAGACATCACCGGTGTGGAACACA 1284
 QY 281 GluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerGln 300
 DB 1285 GAGAGGCGCTATGAGAGTGCAGAGAGTGTGCTGCGGATGAGAGTGCAGCTCCAG 1344
 QY 301 LeuGlnValGlnArgSerLeuThrPheAlaThrGluAspGlnLysIleTyrglyIleThr 320
 DB 1345 CTCACAGTCCAGAGATCCCTGTGACAGCCAGCAGAGACAGAAATCTACATCTACACC 1404
 QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
 DB 1405 CTCAGAGGATGTGCGCTTAAACACACCCCAACAGCGCTGTGATCTCCAGCTGCGAC 1464
 QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrglyLeuValLeuAlaGlyLeu 360
 DB 1465 ACCTGCTTTTGGCGCTGCTGTATTAATAAGAAATCTTACCTGCTTACGCGGCGCTC 1524
 QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 DB 1525 GCCGATGGGCTGTGTGCTGTGTTTCCCGTGGGCGGACCCCAAGAGACAGCTGCTCC 1584

QY 241 PheMetYrGluLeuCySGlyLysGlnThrAlaPhePheSerSerGlnGlnGlu 260
 DB 4844 TTCAATGATGAACGTGTGCTGAGGAGCAGACAGCTTCTTCTCAATCCAGGCGCAGG 4903
 QY 261 TyrThrValValPheThrAspGlyLysGlnSerArgAsnTyrThrValAlaAsnThr 280
 DB 4904 TACACGGTGTGTTTGGGATGGAAAGAGAGATCCAGAACTACACGGTGGTGAACAA 4963
 QY 281 GluYrGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
 DB 4964 GAGAAAGGCGCTCATGAGATGACAGAGATGCTGCTCCCTGGGATGAAGGTGAGCTCCAG 5023
 QY 301 LeuGlnValGlnArgSerLeuThrAlaThrGluAspGlnLysIleTyrIleTyrThr 320
 DB 5024 CTCACAGGTCCAGAGATCCCTGTGAGACAGCCACCGAG----- 5059
 QY 321 LeuYrGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaVal 340
 DB 5059 ----- 5059
 QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
 DB 5060 -----AATTCCTACCTGCTTACGGGCGCTC 5086
 QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 DB 5087 GCCGATGGGCTGTGTGCTGTCTTCCCGGTGGCGGAGCCCAAGAGACAGCTCTCC 5146
 QY 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
 DB 5147 TACCTGTGCTCACACACACACACAGTCCAGATTCAGATGCGGATGAAGACGACAG 5206
 QY 401 GlnAsnProTyrProValLysAlaMetGluValAlaAsnSerGlySerGluValTyrTyr 420
 DB 5207 CAGAAACCTTACCCAGTGAAGCCATGAGGTGCTCAAGCGGCTCTAGGTCTGTAC 5266
 QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
 DB 5267 ACCAATGGGCGCGGCTCTGTCTGATCGACTGCTCCCTGAGATCGACAGCGCGCTG 5326
 QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnArgGly 460
 DB 5327 GAGCCCTTACATGGCCCTCCATGTTACGTACGTGCTGACGCTTGAAGGCAAGAGG 5386
 QY 461 GluGluValValTyrPysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 DB 5387 GAGGAGGTGCTGTGCTGATGATGACAGGCCAACTCTTGTGATGTACACATCTCAC 5446
 QY 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
 DB 5447 ACCTACACACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5506
 QY 501 ProValArgProLeuAspThrGlnProProAlaAlaSerHisThrAlaAsnProLysVal 520
 DB 5507 CCGGTCGCGGCTTGGACACGGAACCCCGGACAGCCACAGCCCAACCAAGGAG 5566
 QY 521 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGluLeuGlnThrGln 540
 DB 5567 CCGTGAAGGGGACTCCATCCGAGACATGATCATCATGATGATGATGATGATGATGATG 5626
 QY 541 IleLeuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSer 560
 DB 5627 ATCTGATGATCCACAGAAATCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5686
 QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
 DB 5687 CCACCCCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5746
 QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
 DB 5747 AGAGTGCCTTCTCCACGACATGCGAGATCAGACATCTCATATGCGCGGCTGCTGCTG 5806

QY 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
 DB 5807 TCCGACAGGTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5866
 QY 621 GlnAlaValLysIleLeuAlaValAlaArgAspLeuIleTyrValProArgArgGlyLysP 640
 DB 5867 CAGGCGGTGAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5926
 QY 641 ValIleValIleGlyLeuGlnLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 DB 5927 GTTATGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5986
 QY 661 LeuYrAlaArgGluLeuThrProHisGlyValIleValAlaValAlaValAlaVal 680
 DB 5987 TTAAAGCCCGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6046
 QY 681 AspThrValValCysThrPheGluAsnGlnValAsnThrGluTyrCysLeuAlaValTyrArg 700
 DB 6047 GACACTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6106
 QY 701 GlyTyrGlyAlaArgGluPheAspIlePheTyrGlnSerTyrGlnGluLeuGlyArgLeu 720
 DB 6107 GCGTGGGCGCGCAGGAGATTCACATTTCTTCAAGTCTTACGAGAGAGCTGGGCGGCTG 6166
 QY 721 GluAlaCysThrArgLysArgArg 728
 DB 6167 GAGGCTTGCACTCGCAAGAGAGAG 6190
 RESULT 7
 US-10-132-382-1
 ; Sequence 1, Application US/10132382
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 7093
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-132-382-1
 Alignment Scores:
 Pred. No.: 5,236-249 Length: 7093
 Score: 3579.50 Matches: 688
 Percent Similarity: 94.51% Conservative: 0
 Best Local Similarity: 94.51% Mismatches: 1
 Query Match: 93.66% Indels: 39
 DB: 40 Gaps: 1
 US-09-836-392-21 (1-728) x US-10-132-382-1 (1-7093)
 QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 20
 DB 4202 ATGTGAGGACACTGGGGGCGACACGATGCCATGAAGAACTTCCGAGTTCCGGCAGAG 4261
 QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleLysIleSerIle 40
 DB 4262 GCCAGATGCTGACGCGCTGAGACACCCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 4321
 QY 41 HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 DB 4322 CACCGCTGTGCTTGGCTTGGAGCTCGGCGGCTCAGACGCTCAACACGCTGCTGCTGCTG 4381
 QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 80
 DB 4382 GAGAAAGCCAGAGATTTCTTATACCCCTGGGACACATGCTCACCAAAAAATATAGCC 4441
 QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysValAsnIleIlePheCysAspLeu 100

Db 4442 TACCAGATCGGCTCGGGCTGGCTTACCTGACACAGAAAAACATCATCTTCTGTGACCTG 4501
 Qy 101 LysSerAspAsnIleLeuValTrrPserLeuAspValLysGluHisIleAsnIleLysLeu 120
 Db 4502 AAGTCGGACAACATTTCTGTGTGTGTCTTACCTGACAGGACATCATCATCACTCACTCA 4561
 Qy 121 SerAspTyrTyrIleSerArgIleSerPheHisGluGluValAlaLeuGlyValGluGlyThr 140
 Db 4562 TCTGACTACGGAGATTTCAGAGCACTCATTCATAGAGGCGCCCTAGAGCGTGGAGGGCACT 4621
 Qy 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet 160
 Db 4622 CCGTGGCACACAGCGCCACAGATACAGGCTCGCATGTATATATATGAGAGATGATATG 4661
 Qy 161 PheSerTyrGlyMetValLleuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlyHis 180
 Db 4682 TTTCCTCATGTGAATGGTGGCTCTACGACTGTCTGACAGACCGCCCTGACCTGGGCGCAC 4741
 Qy 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLleuGlyGlnPro 200
 Db 4742 CACCACTCCAGATTGCCAGAGACCTGTCCAGAGGACATCCGCGGTCTGGGCGAGCGG 4801
 Qy 201 GlnGluValGlnPheArgArgIleGlnAlaLeuMetMetGluCysTrrPaspThrLysPro 220
 Db 4802 GAGAGATGAGATGCCGCGCACGCTGACAGCGCTCATGATGAGTGTGGGACACTAAGCCA 4861
 Qy 221 GlnLysArgProLeuAlaLeuSerValLysSerGlnMetLysAspProThrPheAlaThr 240
 Db 4862 GAGAGAGCGACCGCTGGCCCTGTCTGGTGTGAGCCAGATGAGAGACCCGACTTTGGCCACC 4921
 Qy 241 PheMetTyrGluLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGlu 260
 Db 4922 TTCTGTATGACTGTGTGTGTGGAGAGACAGACCTTCTTCTCATCCAGGCGGACAGAG 4981
 Qy 261 TyrThrValValPheTrrPaspGlyLysGluLeuSerArgAsnTyrThrValValAsnThr 280
 Db 4982 TACACCGTGCTTTGGGAGAGAAAGAGAGTCCAGAGACTACAGCTGGTGTGACACA 5041
 Qy 281 GlnLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
 Db 5042 GAGAGAGCGCTCATGAGAGTGCAGAGATGTGTGCTGCTGGAGATGAGAGTGTGAGCCAG 5101
 Qy 301 LeuGlnValGlnArgSerLeuTrrPthrAlaThrGluAspGlnLysIleTyrIleTyrThr 320
 Db 5102 CTCCAGGTCCAGATTCCTTGTGACACCCAGCCAG----- 5137
 Qy 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
 Db 5137 ----- 5137
 Qy 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
 Db 5138 -----AATTCCTACCTGGTGTCTTGGCGGCGCTC 5164
 Qy 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 Db 5165 GCCATGTGGCTGTGTGTGTGTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5224
 Qy 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
 Db 5225 TACCTGTGCTCACACAGAGCCAGACAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 5284
 Qy 401 GlnAsnProTyrProValLysAlaMetGluValValAlaAsnSerGlySerGluValTrrPyr 420
 Db 5285 CAGAACCCCTACCCAGATGAGAGCCATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5344
 Qy 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
 Db 5345 AGCAATGGGCGGCGCTCTGTATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 5404
 Qy 441 GluProTyrMetAlaProSerMetValThrSerValLysSerSerGlyGlyValArgGly 460
 Db 5405 GAGCCCTACAGAGCCCTCATGT 5464

Qy 461 GlnGluValValTrrPysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 Db 5465 GAGAGAGTCTGT 5524
 Qy 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
 Db 5525 ACCCTACAGCTGT 5584
 Qy 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
 Db 5585 CCGTGGCGGCGCTTGAGACAGAGAACCCCGGACAGCCAGCCAGACAGGCGCAACCCAAAGGTG 5644
 Qy 521 ProGlnLysAspSerIleAlaAspValSerIleMetTyrSerGlnGluGluGlyThrGln 540
 Db 5645 CCGTGGAGGAGCTCCATCGGAGAGT 5704
 Qy 541 IleLeuIleHisGlnIleGlnSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSer 560
 Db 5705 ATCTGTATCCACAGAGATCATCTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 5764
 Qy 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerProAlaAspSerSer 580
 Db 5765 CCACCCCGGACAGCTGCGCAGAGTCCCTCAAGCTCCCTCCAGCTCCCGCAGCAAGTTCCTCC 5824
 Qy 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
 Db 5825 AGTGTGCTTCTTCACCGACTGCGAGACTGAGACTGAGACTGAGACTGAGACTGAGACTGAGACTG 5884
 Qy 601 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
 Db 5885 TCCGACAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5944
 Qy 621 GlnAlaValLysIleLeuAlaValAlaArgAspLeuIleTrrPValProArgArgGlyLysP 640
 Db 5945 CAGGCGCTGAAGATCTTCGCGCTGAGAGACTCATTTGGTGGTCCCAAGCGCGGGGAGAT 6004
 Qy 641 ValIleValIleGlyLeuGluLysAspSerGlnAlaGlnArgValIleAlaVal 660
 Db 6005 GTTATGCTCATGCTGCGGAGAGATGTGGCGCCAGCGGGGCGAGATGATGATGATGATGATGATG 6064
 Qy 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 680
 Db 6065 TTAAAGCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6124
 Qy 681 AspThrValValCysThrPheGlnAsnGluAsnThrGluTrrPysLeuAlaValTrrArg 700
 Db 6125 GACACTGT 6184
 Qy 701 GlyTrrPglAlaArgGluPheAspIlePheTyrGlnSerTyrGluGluValArgLeu 720
 Db 6185 GGTGTGGGCGGCGAGGAGTGTGACATTTTCTACCAAGTCTACGAGAGTGTGGGCGGCTG 6244
 Qy 721 GlnAlaCysThrArgLysArgArg 728
 Db 6245 GAGGCTTGACACTCCAGAGAGAG 6268

RESULT 8
 US-10-132-382-7
 : Sequence 7, Application US/10132382
 : GENERAL INFORMATION:
 : APPLICANT: WEISS, BERTRAM
 : TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 : FILE REFERENCE: SCH-1811
 : CURRENT APPLICATION NUMBER: US/10/132,382
 : CURRENT FILING DATE: 2002-04-26
 : NUMBER OF SEQ. ID NOS.: 26
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ. ID NO. 7
 : LENGTH: 7229
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-10-132-382-7

Alignment Scores:

Pred. No.:	5 35e-249	Length:	72239
Score:	3779.50	Matches:	688
Percent Similarity:	94.51%	Conservative:	0
Best Local Similarity:	94.51%	Mismatches:	1
Query Match:	93.66%	Indels:	39
DB:	40	Gaps:	1

US-09-836-392-21 (1-728) x US-10-132-382-7 (1-7229)

QY	1	MeeuNRGHSLeuAValaIzTRPAspAlaMeTyrAspPheSerGlnPheArgGlnGlu	20
Db	4338	ATGTGTGAGGACCTGGCGGGCCACCAAGTGCATGAAAGAACTCTCCGAGTCCGGACGAGG	4397
QY	21	AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle	40
Db	4398	GCCAGATGCTGCACGGCGCTGCAGCAACCCCTGATCGTGGCTCATGGCATCAGATC	4457
QY	41	HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer	60
Db	4458	CACCGGCTGCTCGCCCTGGAGAGTCGGCCCTCAGAGGCTCAACACCCGTGTCCTC	4517
QY	61	GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnIysIleAla	80
Db	4518	GAGAAACCCAGAGATTTCTTCTTATACCCCTGGACACATGCTCACCCAAAAATAGCC	4577
QY	81	TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu	100
Db	4578	TACCAGATCGCTCGGGCCCTGGCCCTACCTGCACAAAGAAAATCATCTTCTGTGACCTG	4637
QY	101	LysSerAspAsnIleLeuValTTPSerLeuAspValLysGluHisIleAsnIleLysLeu	120
Db	4638	AAGTCGACACACATCTGGTGGTGCTCCCTTGACGTCAAGAGACACATCAACTCAAGCTA	4697
QY	121	SerAspTyrGlyIleSerArgGlnSerPheHisGluGlyAlaLeuGlyAlaLysGlyThr	140
Db	4698	TCTACATCGGGATTTGGAGGACGATCATTCAGAGAGGGCCCTAGCGCTGGAGGGCACT	4757
QY	141	ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet	160
Db	4758	CCTGCTACACAGGGCCCAAGATCAGAGCCCTCGCATGTATGATGAGAAAGGTAAATATAG	4817
QY	161	PheSerTyrGlyMetValIleTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlyHis	180
Db	4818	TTTCCTCATGGAAATGGTGGCTCTACAGATTGCTCTCAGGACAGGCCCTCAGCTGGGCCAC	4877
QY	181	HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuAlaGlnPro	200
Db	4878	CACCAAGCTCCAGATTGGCCAAAGACATGTCCAAAGGAGCATCGCCGGGTCTTGGGGCAGCG	4937
QY	201	GluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGluCysTrpAspPheLysPro	220
Db	4938	GAGAAATGTCAGTTCGGCGCACATCGACAGGGCCCTCATGATGAGAGTCGGGACACTAAGCA	4997
QY	221	GluLysArgProLeuAlaAlaLeuSerValIleSerGlyMetLysAspProThrPheAlaThr	240
Db	4998	GAGAAAGCAGCGCTGGCGCTGCTGGTGGTACGCCAATAGAACCCGACTTTTGGCACCC	5057
QY	241	PheMetTyrGluLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnIleGlnGlu	260
Db	5058	TTTCATGTATGAACATGCTGCTGGGAGCAACAGCTTCTTCCATCCACGAGGCCAGAGAG	5117
QY	261	TyrThrValValPheTTPAspGlyLysGlnGluSerArgAsnTyrThrValValAsnThr	280
Db	5118	TTCACCGGGGTGTGGAGTGGAAAAGAAAGATCCAGAAACTACAGGGTGGGAACACA	5177
QY	281	GluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln	300
Db	5178	GAGAAAGGCCCTCATGGAGGTGCAGAGGATGTGCTGCCCTGGATGAAAGGTGAGCTGCCAG	5237
QY	301	LeuGlnValGlnArgSerLeuTTPThrAlaThrGluAspGlnLysIleTyrIleTyrThr	320

Db	5238	CTCAGAGTCCAGAAATCCCTGTGGACAGCACGCCAGG	-----	5273
Qy	321	LeuylsGIYmetCysProleuaSnrThrProGlnGlnAlaLeuaSpThrProAlaValAl		340
Db	5273	-----		5273
Qy	341	ThrcysPheLeuaAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaLys		360
Db	5274	-----AAATCCACTGCTGTCTTACGGGGCTC		5300
Qy	361	AlaAspGIYLeuValAlaValPheProValValArgGIYThrProLysAspSerCysSer		380
Db	5301	GCCGATGGGGCTTGGGGCTGTGTTCCTCGTGTGGGGACCCCCAAGACAGCTGTCTCC		5360
Qy	381	TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg		400
Db	5361	TACCTGTCTCAGACACAGCCACAGGTCCAACTTCAGCATGTGGCGATGAAAGCCACGG		5420
Qy	401	GlnAsnProYrProValLysAlaMetGluValAlaAsnSerylSerylValAlaTrpTyr		420
Db	5421	CAGAACCCCTACCCAGAGAAAGGCCATGAGAGTGTGTCAACAGCGCTCTGACTGTGATC		5480
Qy	421	SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgArgLeu		440
Db	5481	AGCAATGGGGCCGGCCCTCTGTGTATCAGATGTGCTCCCTGTGAGATGTGGAGGGGGTGG		5540
Qy	441	GluProYrMetAlaProSerMetValThrSerValAlaCysSerSerGluValArgGly		460
Db	5541	GAGCCCTCAATAGGGCCCCCTCCATGTGTACTGATCAGTGTGTGGACGCTCTGAGGCGAAGG		5600
Qy	461	GluGluValValTrpCysLeuaAspAspLysAlaAsnSerLeuValMetYrHisSerThr		480
Db	5601	GAGGAGGTCTGTGTGCTGCTGATGATGACAAAGCCAACTCTGTGATGTACACACTCCAC		5660
Qy	481	ThrYrGlnLeuCysAlaArgYrPheCysGlyValProSerProleuaArgPheThr		500
Db	5661	ACCTACACAGCTGTGTGCCGGGTACTTGTGTGGGGTCCCAAGCCCTCGAGGACATGTT		5720
Qy	501	ProValArgProLeuaAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal		520
Db	5721	CCCGTGGCGCCCTTGAGACAGGAACCCCGGGACCGACGACACAGGCCCAACCAAGGTG		5780
Qy	521	ProGluGlyAspSerIleAlaAspValSerIleMetYrSerGluLeuGlnIleThrGln		540
Db	5781	CCTGAGGGGGACTCATACGCGGAGAGTGTACATCATGTACAGTAGAGAGCTGGGACAGAG		5840
Qy	541	IleLeuIleHisGlnIleuserLeuThrAspTyrCysSerMetSerSerYrSerSerSer		560
Db	5841	ATCCGATGCCACCGAGGATCATCTACAGTACTGTCTCCATGTCTCTCATCTCATCC		5900
Qy	561	ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProLysSerSerSer		580
Db	5901	CCACCCGCCACAGCTGCCAGGTCCCTCAAGGCTCCCAAGCTCCCAAGCTCCCTTCC		5960
Qy	581	SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla		600
Db	5961	AGTGTGCTTCTCTCCACGACTCGGAGAGACTCAGACATGCTACATACGCCCGCTGTCC		6020
Qy	601	SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu		620
Db	6021	TCCGACAGGTGTGACATGACATGATACCCCATGAGCGGGAGACCTTCAAGCAGCACTGG		6080
Qy	621	GlnAlaValLysIleLeuAlaValArgAspLeuIleTrpValProArgArgGlyLysAsp		640
Db	6081	CAGGCCGTAAGATCTCTGCCGTGAGAGACTCATTTGGTCTCCCAAGCGCGGTGGAGAT		6140
Qy	641	ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyLysValIleAlaVal		660
Db	6141	GTTATCTGTCATTTGGCTGGAGAAAGATTCTGGCGCCACAGCGGGCGAGTCAATTGCTGCTC		6200
Qy	661	LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys		680
Db	6201	TTTAAAGCCAGAGCTGACTCCGCAATGAGGGTCTGTGTGATGTCTCCCTGTGTGCAAG		6260

QY 681 AsphthValValCysThrPheGluAsnGluAsnThrGluTrpCysLeuAlaValTTrpArg 700
 |||||
 Db 6261 GACAGCTTGTGTGACCTTGAATAAATGAAACAGAGGTGCTCCGGCGTGTGAGG 6320
 |||||
 QY 701 GATrPValAlaArgGluPheAspIlePheTyrGlnSerTyrGlnGluLeuArgLeu 720
 |||||
 Db 6321 GGCTGGGGCCGACGAGGAGTTCGACATTTTCTACAGTCTTACGAGGAGTGGCGGCTG 6380
 |||||
 QY 721 GluAlaCysThrArgIleArgArg 728
 |||||
 Db 6381 GAGGCTTGACCTGCGCAAGAGAGG 6404
 |||||

RESULT 9

US-10-132-382-3
 ; Sequence 3, Application US/10132382
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 7307
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-132-382-3

Alignment Scores:

Pred. No.: 5,42e-249 Length: 7307
 Score: 3579.50 Matches: 688
 Percent Similarity: 94.51% Conservative: 0
 Best Local Similarity: 94.51% Mismatches: 1
 Query Match: 93.66% Indels: 39
 DB: 40 Gaps: 1

US-09-836-392-21 (1-728) x US-10-132-382-3 (1-7307)

QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheSerGluPheArgGln 20
 |||||
 Db 4416 ATCTGAGGACCTGGCGGCGACGATGCATGAAGAATCTTCGAGTCCGGCAGAG 4475
 |||||
 QY 21 AlSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
 |||||
 Db 4476 GCCAGCATGCTGCACGCGCTGCAGACACCTGATCTGCGGCGCATCGCATCAGCATC 4535
 |||||
 QY 41 HisProLeuGlyPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 |||||
 Db 4536 CACCCGCTGTGCTTGGCTGGAGCTGCGCCGCTCAGACGCTCAACACCGCTGCTGCC 4595
 |||||
 QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 80
 |||||
 Db 4596 GAAAGGCCAGAGATTTCTTATACCCCTGGGACACATGCTCACCCAAAAAATAGCC 4655
 |||||
 QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
 |||||
 Db 4656 TACCAGATCCCTCGGGCTGCGCTACCTCACAAGAAAAACATCATCTCTGTGACCTG 4715
 |||||
 QY 101 LysSerAspAsnIleLeuValIleTrpSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
 |||||
 Db 4716 AACTCGACAAACATTCGTGGGTGCTGCTTGAAGTCAAGAGCAACATCAATCACTA 4775
 |||||
 QY 121 SerAspTyrGlyIleSerArgGlnSerPheHisGlnGluValAlaLeuGluValGlnGlyThr 140
 |||||
 Db 4776 TCGAGCTAGAGGATTCGAGGAGTATTCATGAGAGGCGCCCTAGGCGTGGAGGCGACT 4835
 |||||
 QY 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrArgGlnLysValAspMet 160
 |||||
 Db 4836 CCGTGGCTACAGGCCCGCAGAGATCAGGCGCTCATGTATATGATGAGAAAGGTGATATG 4895
 |||||
 QY 161 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyLysArgProAlaLeuGlnHis 180
 |||||

Db 4896 TTCCTATGAGATGGTGGCTTACGAGTGTCTGCAGAGACGCGCTGACCTGGGCGAC 4955
 |||||
 QY 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnPro 200
 |||||
 Db 4956 CACAGCTCCAGANTGGCAAGAGAGCTCCAGAGGATCCGCCGCTTCCGGGCGACCGC 5015
 |||||
 QY 201 GluGluValGlnPheArgArgLeuGlnAlaLeuMetLeuIleCysTrpAspThrLysPro 220
 |||||
 Db 5016 GAGAAAGTGCAGTCCCGCCAGCTGCAGCGCTCATGTATGAGAGTCTGGGACATAAGCCA 5075
 |||||
 QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
 |||||
 Db 5076 GAGAGGACCGGCTGGCGCTGCTGCTGATGACCAAGTGAAGACCCGACTTTGGCCACC 5135
 |||||
 QY 241 PheMetTyrGlnLeuGlyCysGlyLysGlnThrAlaPhePheSerSerGlnGlnGlu 260
 |||||
 Db 5136 TTGATGATGACAGCTGCTGCTGGAGAACACAGACCTCTTCTCATCCAGGCGCAGAG 5195
 |||||
 QY 261 TyrThrValValPheTrpAspGlyLysGlnGluSerArgAsnTyrThrValAlaAsnThr 280
 |||||
 Db 5196 TACACCGTGTGTTTGGATGGAAGAGAGAGTCCAGAACTACACGCTGGTGAACACA 5255
 |||||
 QY 281 GluLysGlyLeuMetGluValGlnArgMetCysSerProGlyMetLysValSerCysGln 300
 |||||
 Db 5256 GAGAGGGCTCATGGAGGTCAGAGAGTGTCTGCTGCGTGGAGTGAAGGTGACTGCGCAG 5315
 |||||
 QY 301 LeuGlnValGlnArgSerLeuTrpThrAlaThrGlnAspGlnLysIleTyrIleTyrThr 320
 |||||
 Db 5316 CTCAGGTCCAGAGATCCCTGCTGAGACGCCAGCAG----- 5331
 |||||
 QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAla 340
 |||||
 Db 5351 ----- 5351
 |||||
 QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
 |||||
 Db 5352 -----AATTCCTACCTGCTGCTTACGGGGCTC 5378
 |||||
 QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 |||||
 Db 5379 GCCGATGGCTGTGGCTGTGTTCCCGTGGTGGGCGACCCCAAGGACAGCTGCTCC 5438
 |||||
 QY 381 TyrLeuGlySerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
 |||||
 Db 5439 TACCTGTGTCACACACAGCCCAACAGGTCCAAAGTTCAGCATCCGGATGAAGCGCAGG 5498
 |||||
 QY 401 GluAspProTyrProValLysAlaMetGlnValValAsnSerGlySerGluValTTrpTyr 420
 |||||
 Db 5499 CAGAACCCCTAACCCAGTGAAGGCCATGAGGTGATCAACAGCGGCTTGAGGTCTGTAC 5558
 |||||
 QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnLysCysArgArgLeu 440
 |||||
 Db 5559 AGCAATGGCGCGGCTCCCTGTTCATGAGATGTGCTCCCTCGAGATCTCAGCGGCTG 5618
 |||||
 QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgGly 460
 |||||
 Db 5619 GAGCCCTACATGGCGCCCTCATGAGTTAAGTCAAGTCTGTGCGAGCTGAGGCGAGAGG 5678
 |||||
 QY 461 GluGluValValTTrpCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 |||||
 Db 5679 GAGAGGTGCTGTGCTGCTGATGACAAAGGCCAACTCTGTGATGATCCACTCCACC 5738
 |||||
 QY 481 ThrTyrGlnLeuGlyAlaArgTyrPheCysGlyValAlaProSerProLeuArgAspMetPhe 500
 |||||
 Db 5739 ACTTACAGAGTGTGTGCGGCTACTTCTGGGGGTGCCAGCCCTCAGAGCAAGTTT 5798
 |||||
 QY 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
 |||||
 Db 5799 CCGTGGCGGCTTGGACAGGAAACCCCGCAGCCAGCCACAGCGCAACCCAAAGGTG 5858
 |||||
 QY 521 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGluLeuGlyThrGln 540
 |||||

Db	5859	CCTGAGGGGGACATCCATTCGGCGAGAGTACATCATCTACAGTAGAGAGAGCTGGGACAGGAG	5918
Oy	541	ITLeuLeuIHsIngInguSerLeuThrAspTryCysSerMetSerSerTrySerSerSer	560
Db	5919	ATCTCGATCCACCGAGGAATACCTACTGATGTAATGCTCATGTCATCTCTCTACCTCCACAC	5978
Oy	561	ProProlArgGlnAlaAlaArgSerProSerSerLeuProSerSerProIAsSerSer	580
Db	5979	CCACCCCGCCAGGCTGCCAGGTCCCTCCCTCAAGCTCTCCCAAGCTCCCGACAGAGTTCTTCC	6038
Oy	561	SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGluAlaAla	600
Db	6039	AGTGGCCTTTCTCCACCCGACTCGAGAGACTGACATGCTCAATAGACCGCGGTCTGCC	6098
Oy	601	SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu	620
Db	6099	TCCACACAGGTCTGACATGACTGACCCCTGAGCGGGGAGAACCTTCACGCCACGACCTG	6158
Oy	621	GlnAlaValLysIleLeuAlaAlaValArgAspLeuIleTrpValProArgArgGlyGlyAsp	640
Db	6159	CAGGCCCGGAAGACCTCTGCCGTCACAGACACTATTGGGTGCCCAAGCGCGGTGGAAAT	6218
Oy	661	ValIleValIleGlyLeuGluLysAspSerGlnAlaGlnArgGlyArgValIleAlaVal	660
Db	6219	GTTATCGCATTTGGCCCTGGAGAAAGATTCCTGGCGGCCACGGGCGCGAGTCAATCTCCGTC	6278
Oy	661	LeuLysAlaAlaArgGluLeuThrProHisGlyValIleValAspAlaAlaValAlaLys	680
Db	6279	TTAAAGCCCGAGAGCTGACTCCGACAGGGGGTCTGTGGATGCTGCCCTGGTGGCAAG	6338
Oy	681	AspPheValValCysThrPheGluAsnGlnAsnThrGlyTrpCysLeuAlaValTrpArg	700
Db	6339	GACACTGTGTGCTCACCTTGAAGATGAAGAACAGAGATGGTGGCTGGCCCTCTCGAGAG	6398
Oy	701	GlyTrpGlyAlaArgGluPheAspIlePheTryGlnSerTryGluGluLeuGlyArgLeu	720
Db	6399	GGCTGGGGCGCCAGAGGAGTTCGACATTTCCTACCACTCTACAGAGAGCTGGCGGGCTG	6458
Oy	721	GluAlaCysThrArgLysArgArg	728
Db	6459	GAGGCTTGCACCTCGCAAGAGAGG	6482

RESULT 10
PCT-US02-05109-456
Sequence 456, Application PC/TUS0205109
GENERAL INFORMATION:
APPLICANT: HySeq Inc.
APPLICANT: Tang, Tom Y.
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Ma, Yundong
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-thong
APPLICANT: Wang, Zhimei
APPLICANT: Ghosh, Malabika
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dунrui
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21772-094-061
CURRENT APPLICATION NUMBER: PCT/US02/05109
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 09/810,173
PRIOR FILING DATE: 2001-03-15
NUMBER OF SEQ ID NOS: 1052
SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 456
; LENGTH: 2839
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (872)..(1834)
PCT-US02-05109-456

Alignment Scores:
Pred. NO.: 7.04e-129 Length: 2839
Score: 1919.00 Matches: 372
Percent Similarity: 94.49% Conservative: 5
Best Local Similarity: 93.23% Mismatches: 17
Query Match: 50.21% Indels: 6
DB: 1 Gaps: 1

US-09-836-392-21 (1-728) x PCT-US02-05109-456 (1-2839)

QY 335 AsphtrProAlaValAlValThrCysPheLeuAlaValProVal-----1le 349
Db 636 GATACACACTTCCAGATCCAGTCCGTGCTCTGTCAGGGTTGGGGCTCTTCACACTGT 695
QY 350 LysLysAnSerTYrLeuValAlaGlyLeuAlaAspGlyLeuValAlaValPhePro 369
Db 656 GCGAGGAAGGTGACTCTGGTCTTACGCGCTTGGCGGAGTGGCTTGTCTGTTCCCG 754
QY 370 ValValArgGlyThrProLYsAspSerCysSerTYrLeuCYsSerHisThrAlaAnArg 389
Db 755 TGTGTGGGGGGGACCCCAAGAGAGAGCTGCTCTACCTGTGTCACACACAGCAACAG 814
QY 380 SerLYsPheSerIleAlaAspGluAspAlaArgGlnAsnProTYrProValLYsAlaMet 409
Db 815 TCCAAGTTCACAGCAATCGCGGATGAAGACGACGGCAGAACCCCTACCCAGTGAAGGCCATG 874
QY 410 GluValValAnSerGlySerGluValThrPTyrSerAnGlyProGlyLeuLeuValIle 429
Db 875 GAGGGGGTCAACACAGGCTCTGAGTCTGTGTATACAGCAATGGGGGCTCTTGTATATC 934
QY 430 AspCysAlaSerLeuGlnIleCYsArgArgLeuGlnProTYrMetAlaProSerMetVal 449
Db 935 GACTGTGCTCTCCCGGAGATCTGACAGGGGGCTTGAGCCCTACATGGCCCCCTCATGGTT 994
QY 450 ThrSerValValCYsSerSerGluGlyArgGlyGluGluValAlaTrpCYsLeuAspAsp 469
Db 995 ACGTACAGTCGTGTCACTCTCGAAGGACAGAGGAGAGAGAGTGTCTGTGCTGGATGAC 1054
QY 470 LysAlaAnSerIleValMetTYrHisSerThrThrTYrGlnLeuCYsAlaArgTYrPhe 489
Db 1055 AAGGCAACTCTCTGGTGGATGTACACTCCACACACTTACCAAGTGTGTCCCGGATCTTC 1114
QY 490 CysGlyValIProSerProLeuArgAspMetPheProValArgProLeuAspThrGluPro 509
Db 1115 TGGGGGGTCCCCACCCCCCTTCAGGAGCATTTTCCCGTGGGGCTTGGACACAGAACCC 1174
QY 510 ProAlaAlaSerHisThrAlaAsnProLYsValProGluGlyAspSerIleAlaAspVal 529
Db 1175 CCGGAGCCACAGCCACAGGCCAACCCAAAGTGTCTGAGGGGAGATCCATCCGGGACGTG 1234
QY 530 SerIleMetTYrSerGluGluLeuGlyTYrGlnIleLeuIleHisGlnLeuSerLeuThr 549
Db 1235 AGCAATCATGTCACTGTGAGGAGCTGGGACCCACATCTGTATCCACAGGAATCATCACTACT 1294
QY 550 AspTYrCYsSerMetSerSerTYrSerSerSerProProArgGlnAlaAlaArgSerPro 569
Db 1295 GACTACTCTCTCAATGCTCTCTTACTCTTCAATCCACACCCCGGAGGCTCCACAGTCCCC 1354
QY 570 SerSerLeuProSerSerProAlaSerSerSerSerValProPheSerHisThrAspGly 589
Db 1355 TCAAGCTCTCCACACTCCCAAGCAAGTTCTTCAAGTGTGCTTCTTCCACAGCATGGCAG 1414
QY 590 AspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgSerGluHisAspLeuThr 609

```

Db	1415	GACTCAGACATGCTACATACAGCCCGGGTGCTGCCTCCGACAGGCTGAGAGTACCTGACC	1474
QY	610	ProMetAspGlyGluThrPheSerGlnHisLeuGlnAlaValLysIleLeuAlaValArg	629
Db	1475	CCCATGGACCGGAGACCTTTCACCCAGCACCTGACAGGCGGTGAAATCTCCCTCCCTGAGA	1534
QY	630	AspLeuIleThrValProArgArgGlyLysAspValIleValIleGlyLeuGluLysAsp	649
Db	1535	GACCTCATTTGGGTGCTCCCGACGGCGCGGTGGAGATGTTATCGTATTTGGCTGGAGAAAGAT	1594
QY	650	SerGlnAlaGlnArgGlyArgValIleAlaValLeuLysAlaArgGluLeuThrProHis	669
Db	1595	TCTGGCGCCCGACGGGGCCGACGATCTGGCCGTCTTTAAAGCCCGAGAGCTGACTCCGAT	1655
QY	670	GlyValLeuValAspAlaAlaValAlaValAlaLysAspThrValValCysThrPheGluAsn	689
Db	1655	GGGGTGCTGGTGGATGCTGCCGCGTGGCAAAAGACACTGTTGTGTGCACCTTGAAAT	1714
QY	690	GluAsnThrGluTrpCysLeuAlaValIleThrArgGlyTTPGlyValaArgGluAspIle	709
Db	1715	GAAMAACACAGAGTGCTGCCCTGGCCGTCTGGAGGGGCTGGGGGCCACAGGAGTTGCACATT	1774
QY	710	PheTrpGlnSerIleTrpGluLeuGlyArgLeuGlnAlaCysThrArgLysArgArg	728
Db	1775	TTCTACCAAGTCTTCACGAGAGAGCTGGGGCCGCTGAGAGGCTTGACCTGCACAAAGGAAGG	1831

D	73	AACTTCAGCATGGGGGATGGAAGACGACGAGGAGAACCCCTAACCCAGAGGAAGCCATGGAG	132
Q	411	ValValaInserGlySerGluValITPTySerAsnGlyProGlyLeuLeuValIleAsp	430
D	133	GTGGTCAACAGCGGCTGTGAGGTGTGGTATACAGCAATGAGGGCGGGCCCTTGTGATCAG	192
Q	431	CysAlaSerLeuGluIleCysArgArgLeuGluProTyrMetAlaProSerMetValThr	450
D	193	TGTGCTCTCCCTGGAGATCTGCAGGGCGGCTGGAGGCCCTTACATGGCCCCCTCATGTGTTACG	253
Q	451	SerValValCysSerSerGluGlyArgGlyLeuValValITPTyCysLeuAspAspLys	470
D	253	TCACTCGGTGTCACCTCTGAGAGGGCAGAGGGAGAGAGGTGCTGCTGGTCCGTGATGACAG	312
Q	471	AlaAsnSerLeuValMetTyrHisSerThrTyrTrpGluLeuCysAlaIleArgTyrPheCys	490
D	313	GCCAACTCCCTGGTGGATGTACCACTCCACCACTCACAGCTGTGTGGCCGGTACTTGTGC	377
Q	491	GlyValProSerProLeuArgAspMetPheProValArgProLeuAspThrGluProPro	510
D	373	GGGCTCCCCAGCCCCCTCAGGAGCATGTTTCCGTGGCGCCCTTGGACAGGAAACCCCCG	433
Q	511	AlaAlaSerHisThrAlaAsnProLysValProGluLysSerIleAlaAspValSer	530
D	433	GCACCCAGCCACAGCGGCAACCCAAAGGTGCTGAGGGGGAGCTCCATCGCGGAGCTGAGC	492
Q	531	IleMetTyrSerGluGluLeuGlyThrGlnIleLeuIleHisGlnGluSerLeuThrAsp	550
D	493	ATCATGTACAGTGTAGGAGCTGGGACGACGAGATCCTGTATCCACAGATCACTCACTGAC	552
Q	551	TyrCysSerMetSerSerTyrSerSerSerProProArgIleAlaIleArgSerProSer	570
D	553	TACGTCTCCACATGCTCTTACTCTCTCATCCACCCACCCGCGCAGGTGGCAGGTCCCTCA	612
Q	571	SerLeuProSerSerProAlaSerSerSerSerValProPheSerThrAspCysGluAsp	590
D	613	AGCCTCCCACTCCCTCCACAGCAAGTCTTCCAGTGTGCTTCTCCACATGATGGAGAGAC	672
Q	591	SerAspMetLeuHisThrProGlyAlaAlaSerAspArgSerGluHisAspLeuThrPro	610
D	673	TCAACATGCTACATACAGCCCGGCTGTGCTCCACAGAGTGTGAGCATGACCTGACCCCC	732
Q	611	MetAspGlyGluThrPheSerGlnHisLeuGlnIleAlaValLysIleLeuAlaValArgAsp	630
D	733	ATGACCGGGGAGACCTTCAGACACACCTCGAGGCCGGGAGATCTTCGCGTAGAGAC	792
Q	631	LeuIleThrValProArgArgGlyGlyAspValIleValIleGlyLeuGluLysAspSer	650
D	793	CTCATTTGGGTCCCAAGCGCGGGGTGGAGATTTTATCGTCATTTGGCTTGGAGAAAGATTCT	853
Q	651	GluAlaGlnArgGlyArgValIleAlaValLeuLysAlaArgGluLeuThrProHisGly	670
D	853	GGCGGCCAGCGGGCGGCGAGTATTGGCCGCTTAAAGCCGAGAGCTGATCCCATGGG	912
Q	671	ValLeuValAspAlaAlaValValAlaLysAspThrValValCys-ThrPheGluAsnGln	690
D	913	GTGCTGGTGATGTGTGCCCCGGAGGACCAAGCACTTGTGTGTGCACCTTTGAAAGA	972
Q	690	wasnThrGluTyrCysLeuAla	697
D	973	AAACACAGAGGGGTGCTTGGC	994
RESULT 13			
US-09-637-890-9034			
Sequence 9034, Application US/09637890			
GENERAL INFORMATION:			
APPLICANT: Gearlog, David P.			
APPLICANT: Shyjan, Andrew W.			
APPLICANT: Holtzman, Douglas A.			
APPLICANT: Galvin, Katherine			
APPLICANT: Culpepper, Janice A.			
APPLICANT: Leidy, Kevin R.			
APPLICANT: Vasicek, Tom			

/ APPLICANT: Macbeth, Kyle J.
 / APPLICANT: Vallejo, Jean-Luc M. G.
 / APPLICANT: Cepeda, Mario
 / APPLICANT: Kingsbury, Gillian A.
 / APPLICANT: Bustfield, Samantha J.
 / TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 / FILE REFERENCE: 1600,1152-001
 / CURRENT APPLICATION NUMBER: US/09/637,890
 / CURRENT FILING DATE: 2000-08-09
 / PRIOR APPLICATION NUMBER: 60/147,939
 / PRIOR FILING DATE: 1999-08-09
 / NUMBER OF SEQ ID NOS: 10217
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 9034
 / LENGTH: 1912
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-09-637-890-9034

Alignment Scores:
 Pred. No.: 1,18e-107 Length: 1912
 Score: 1624.00 Matches: 311
 Percent Similarity: 99.68% Conservative: 1
 Best Local Similarity: 99.36% Mismatches: 1
 Query Match: 42.49% Indels: 0
 DB: 24 Gaps: 0

US-09-836-392-21 (1-728) x US-09-637-890-9034 (1-1912)

QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLeuAspPheSerGluPheArgGlnu 20
 DB 972 ATGCTAGGACCTGGGGGCGACCGCATGACATGACATGCTCCAGTTCGGCGAGAG 1031
 QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
 DB 1032 GCCAGCATGCTGCACCGCTGCAGCAGCCCTGCATGCTGCCTCATCGGCATCAGCATC 1091
 QY 41 HisProLeuCysPheAlaLeuGlnuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 DB 1092 CACCCCTGCTGCTCCCTGGAGCTGGCGCCCTCAGCAGCCCTCAACCCGCTGTCTCC 1151
 QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnuIleAla 80
 DB 1152 GAGAACGCCAGAGATCTTCTTATACCCCTGGACACTGCTCAACCCAAAATAAGCC 1211
 QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
 DB 1212 TACCAATGCGCTCGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1271
 QY 101 LysSerAspAsnIleLeuAlaTyrSerLeuAspValLysGlnHisLysAsnIleLysLeu 120
 DB 1272 AGTGTGGACAACTTGT 1331
 QY 121 SerAspTyrGlyIleSerArgGlnSerPheHisGlyIleAlaLeuGlyValGlnuIleThr 140
 DB 1332 TGTGACTAGGAGGATTCGAGGCGATTCATCCATGAGGCGCCCTAGCGCGGAGGCACT 1391
 QY 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet 160
 DB 1392 CCGGCTACAGGCGCCCAAGATCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1451
 QY 161 PheSerTyrGlyMetValLeuTyrGlnuLeuSerGlyLysArgProAlaLeuGlyHis 180
 DB 1452 TTCTCTATAGGAATGTGTCTTACGAGTGTCTTACGAGAGAGCGCCCTGACCTGGCCAC 1511
 QY 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGlnPro 200
 DB 1512 CACCACTCAGATTCGCAAGAGCTGCTCAAGGCGATCCGCCGCTTCGGGCGAGCGG 1571
 QY 201 GlnGlnValGlnPheArgArgLeuGlnAlaLeuMetMetGluCysTyrAspThrLysPro 220
 DB 1572 GAGGAGTGCAGTTCGCGGAGCTGACGAGGCGCTCATGTAGTAGTGTGGAGCACTAAGCCA 1631

QY 221 GlnLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
 DB 1632 GAGAACGCCAGCACTGCGCTGTGCTGTGAGCCCAATGAAGACCGCACTTGTGCGACC 1691
 QY 241 PheMetTyrGlnLeuCysGlyLysGlnThrAlaPheSerSerGlnuIleGlnu 260
 DB 1692 TCTCATATGACTGCTGCTGTGAGGAGCAGACAGCTTCTTCTCATCCAGGCGCAGAG 1751
 QY 261 TyrThrValAlaPheTyrAspGlyLysGlnuLeuSerArgAsnTyrThrValAlaSerThr 280
 DB 1752 TACACTGT 1811
 QY 281 GlnLysGlyLeuMetLysValGlnArgMetCysCysProGlyMetLysValSerGln 300
 DB 1812 GAGAAAGGCGCTCATGAGATGTCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1871
 QY 301 LeuGlnValGlnArgSerLeuTyrPheAlaThrGluAsp 313
 DB 1872 CTCAGGTCCAGAGATCCCTGTGACAGCCACCGAGAGAC 1910

RESULT 14
 US-09-644-869-9514
 / Sequence 9514, Application US/09644869
 / GENERAL INFORMATION:

/ APPLICANT: Shyjan, Andrew W.
 / APPLICANT: McCarthy, Sean A.
 / APPLICANT: Holtzman, Douglas A.
 / APPLICANT: Monahan, John
 / APPLICANT: Richardson, Jennifer
 / TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 / FILE REFERENCE: 1600,1182-001
 / CURRENT APPLICATION NUMBER: US/09/644,869
 / CURRENT FILING DATE: 2000-08-28
 / PRIOR APPLICATION NUMBER: 60/151,062
 / PRIOR FILING DATE: 1999-08-27
 / NUMBER OF SEQ ID NOS: 9708
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 9514
 / LENGTH: 1912
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-09-644-869-9514

Alignment Scores:
 Pred. No.: 1,18e-107 Length: 1912
 Score: 1624.00 Matches: 311
 Percent Similarity: 99.68% Conservative: 1
 Best Local Similarity: 99.36% Mismatches: 1
 Query Match: 42.49% Indels: 0
 DB: 25 Gaps: 0

US-09-836-392-21 (1-728) x US-09-644-869-9514 (1-1912)

QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLeuAspPheSerGluPheArgGlnu 20
 DB 972 ATGCTAGGACCTGGGGGCGACCGCATGACATGACATGCTCCAGTTCGGCGAGAG 1031
 QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
 DB 1032 GCCAGCATGCTGCACCGCTGCAGCAGCCCTGCATGCTGCCTCATCGGCATCAGCATC 1091
 QY 41 HisProLeuCysPheAlaLeuGlnuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 DB 1092 CACCCCTGCTGCTCCCTGGAGCTGGCGCCCTCAGCAGCCCTCAACCCGCTGTCTCC 1151
 QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnuIleAla 80
 DB 1152 GAGAACGCCAGAGATCTTCTTATACCCCTGGAGCACTGCTCAACCCAAAATAAGCC 1211
 QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100


```
Db 14584 GCCAGAGCTTAAGTGCACACAAAGAAAGCCAGTCACTGTGACTCAGTTCCAC 14643
QY 635 ----- 635
Db 14644 TGAGAAATTCCTCTCTGCTTTGAAGTCTGCCCTCCCATTAATTATCCCTTAGCAG 14703
QY 635 ----- 635
Db 14704 AAATGCAATGTCTGTGCTCCCTCCGACACTTTTGACACACGGCTTCACCTCCCGGA 14763
QY 636 -----ArgArgGlyGly 639
Db 14764 GAGCAGGCTACCGACAAATCGCTTCCCTCTGGGTTTGTGCTTAGCGCGGTGGA 14823
QY 640 AspValIleValIleGlyLeuGlyLysAspSerGluIleGlnArgGlyArgValIleAla 659
Db 14824 GATGTATCTCTCATTTGGCCCTGGAGAAAGATTCTGGCCGCCCAAGCGGGCGAGTCATTGCC 14883
QY 660 ValLeuLysAlaArgGluLeuThrProHisGly 670
Db 14884 GTCTAAAGCCCGAGAGCTGACTCCGCAATGGG 14916
```

Search completed: April 15, 2003, 09:44:46
Job time : 2/55 secs

8	311	8.1	3138	1	PCT-US03-004216-55	Sequence 55, Appl
9	311	8.1	3138	1	US-10-365-0027-55	Sequence 55, Appl
10	311	8.1	3138	11	US-60-426-864-1	Sequence 11, Appl
11	307.5	7.8	2865	9	US-10-263-929-83	Sequence 83, Appl
12	296.5	7.8	3435	1	PCT-US002-24567-86	Sequence 86, Appl
13	296.5	7.7	3435	7	US-10-210-120-66	Sequence 86, Appl
14	294.5	7.7	3143	7	US-09-949-016-1327	Sequence 1327, Ap
15	294.5	7.7	3143	10	US-60-455-444-1697	Sequence 1697, Ap
16	294.5	7.7	3143	11	US-60-443-566-682	Sequence 682, App
17	294.5	7.7	3143	11	US-60-452-680-6705	Sequence 6705, Ap
18	290.5	7.6	3009	9	US-10-263-929-91	Sequence 91, Appl
19	278	7.3	3619	8	US-10-115-678-15	Sequence 15, Appl
20	278	7.3	3717	8	US-10-210-130-129	Sequence 129, App
21	277.5	7.3	2801	9	US-10-144-771-3453	Sequence 3453, Ap
22	277	7.2	1985	6	US-09-724-676-19817	Sequence 1817, A
23	277	7.2	1985	6	US-09-724-676A-19817	Sequence 1817, A
24	273.5	7.2	4017	1	PCT-US002-39126-55	Sequence 55, Appl
25	272	7.1	3141	9	US-10-258-106-27	Sequence 27, Appl
26	270.5	7.1	3538	8	US-10-365-892-36	Sequence 36, Appl
27	270.5	7.1	3538	9	US-10-288-798-36	Sequence 36, Appl
28	269	7.0	3396	6	US-09-724-676-19818	Sequence 1818, A
29	269	7.0	3396	6	US-10-361-942-1504	Sequence 1504, Ap
30	268.5	7.0	1317	9	US-10-144-771-9827	Sequence 9827, Ap
31	268.5	7.0	2553	9	US-10-263-929-90	Sequence 90, Appl
32	268.5	7.0	2553	9	US-10-210-130-31	Sequence 31, Appl
33	266.5	7.0	3057	8	PCT-US002-34037-3	Sequence 3, Appl
34	266.5	7.0	3111	1	US-10-263-929-87	Sequence 87, Appl
35	266.5	7.0	3111	9	PCT-US002-34037-1	Sequence 1, Appl
36	266.5	7.0	5549	1	PCT-US003-02588-23	Sequence 23, Appl
37	266.5	7.0	5549	9	US-10-354-358-23	Sequence 23, Appl
38	266	7.0	2250	11	US-60-452-680-10923	Sequence 10923, A
39	266	7.0	2250	11	US-10-310-154-156	Sequence 156, App
40	265	6.9	2959	7	US-09-837-604A-44609	Sequence 44609, A
41	264	6.9	692	7	US-60-443-566-2179	Sequence 183, App
42	263	6.9	1955	10	US-60-443-566-2179	Sequence 183, App
43	263	6.9	1955	11	US-60-455-444-183	Sequence 183, App
44	262.5	6.9	4450	11	US-60-455-444-183	Sequence 183, App
45	262.5	6.9	4450	11	US-60-455-444-183	Sequence 183, App

```

: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No: LG:337452.25:2002ZAN18
: FEATURES:
: NAME/KEY: unsure
: LOCATION: (1) ... (4120)
: OTHER INFORMATION: a, t, c, g, or other
: PCT-US03-01363-66

```

Pred. No.:	6,35e-263	Length:	4120
Score:	3378.50	Matches:	690
Percent Similarity:	99.13%	Conserved:	0
Best Local Similarity:	99.13%	Mismatches:	1
Query Match:	99.65%	Indels:	50
DB:	1	Gaps:	1

US-09-836-392-21 (1-728) x PCT-US03-01363-66 (1-4120)

D	3	ATCAGCATCACCCGGCTCTGCTGGCCCTGGAGGTGCGCGGCTCAGACACCTCAACACC	62
QY	38	ILeserILeHsPProLeuCySPhaLeaLLeuGLuLeuLaProLeuSerSerLeuAsnThr	57
QY	58	ValLeuSerGIuSnaAlaAaSPSerSerPheILeProLeuGLyHsMetLeuThrGln	77
D	63	GTGTGTCGAGAAACGGCAAGATCTTCTTATACCCCTGGAGACATGCTCACCCAA	122
QY	78	LysILeAlaTYrGlnILeAlaSerGlyLeuAlaTYrLeuHsILysLysAsnILeIlePhe	97
D	123	AAAAATAGCCTTACCAGATCGCTGGCCCTGGCTTACCTGTGCACAAAGAAAACATCATCTTC	182
QY	98	CysAspLeuLysSerAspAsnILeLeuValTRPSerLeuAspValLysGluHsILeAsn	117
D	183	TGTGACCTGAAGTGGAGCAACATCTGTGTGTGGTCCTTGACGTCAAGAGACATCAAC	242
QY	118	ILeLysLeuSerAspTYrGlyILeSerArgGlnSerPheHsIsgluGlyAlaLeuGlyVal	137
D	243	ATCAAGCTACTGTGACTCCGGGATTTCAGAGCAGCATTCATCATAGGGCGCCCTTAGGGCTG	302
QY	138	GluGlyThrProGlyTYrGlnAlaProGluILeArgProArgIleValLYrAspGluLys	157
D	303	GAGGGCACTCTCTGGCTTCCAGGCCCCAGAAATCAGGGCTTGCAATGTATATGATGAAAG	362
QY	158	ValAspMetPheSerTYrGlyMetValLeuTYrGluLeuLeuSerGlyGlnArgProAla	177
D	363	GTAGATATGTCTTCCTATGGAATGAGTGCTCTACAGAGTGTGTGCAGGACAGCCGCCCTGCA	422
QY	178	LeuGlyIshHsIsglnLeuGlnILeAlaLysLysLeuSerLysGlyIleArgProValLeu	197
D	423	CTGGGCCACACCAACGCTCCAGATTTGCCAAAGCTGTGCCAAAGGGCATCCCGGGTCTG	482
QY	198	GlyGlnProGluGluValGlnPheArgArgLeuGlnAlaLeuMetGluCysTYrAsp	217
D	483	GCGCAGCCGAGGAAGCGCATGTCGGCGACTGCAAGCGCTCATGATGAGTCTGGAC	542
QY	218	ThrLysProGluLysArgProLeuAlaLeuSerValValAspGlnMetLysAspProThr	237
D	543	ACTAAGCCAAAGAAAGCAGCCGTGGCCCTGTGGTGTGACCAAGATGAAGAACCCGACT	602
QY	238	PheAlaThrPheMetTYrGluLeuCySArgGlyLysGlnThrAlaPhePheSerGln	257
D	603	TTTGCCACCTTCATGTAATGAACTGTGTGGGAAAGCAACACCTTCTTCATCCAC	662
QY	258	GlyGlnGluTYrThrValValPheTrpAspGlyLysGluLysSerArgAsnTYrThrVal	277
D	663	GCGCAGAGATACACTGTGTGTGGTGGGAGGAAAAAGAGATGCCAGAACTACAGGTC	722
QY	278	ValAsnThrGluLysGlyLeuMetGluValGlnArgMetCysAspProGlyMetLysVal	297
D	723	GTCGAACACAGAAAGAGGCGCTCATGAGAGTCAAGAGATGTGTCGCCCTGGAGTGAAGTGC	782
QY	298	SerCysGlnLeuGlnValGlnArgSerLeuTYrThrAlaThrGluAspGluLysIleTYr	317
D	783	AGCTGCACGCTTCAGAGGTCAGAAATCCCTGTGACAGCCACCGAGAGCCAGAAAATCTTAC	842

OY	318	IleYrThrIleuLysGSLMeTcYsProLeuAsnThrProGlnGlnAlaLeuAspPhePro	337
Db	843	ATCTACACCCCTCAAGGGCAATGTGGCCCTTAAACACACCCCAAGGCCCTTGATCTCTCA	902
OY	338	AlaValValThrCysPheLeuAlaValProValIleLeuLysAsnSerTyrIleuValLeu	357
Db	-903	GCTGCGTACACTCTCTTCTGGCCGCGCTGTTATTTAAAGAATCTCACTTGGTCTTA	962
OY	358	AlaGlyLeuAlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAsp	377
Db	963	GCGGGCCCTCGCCGATGGGGCTGTGGCTGTCTCCGTGGTGGGGGCACCCCAAGGAC	1022
OY	378	SerCysSerTyrIleuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlu	397
Db	1023	AGCTGCTCTTACCTGTCTCTACACACAGCCAAAGGTCCTCAAGTTCAGCATCCGGATGA	1088
OY	398	AspAlaArgGlnAsnProTyrProValLysAlaMetGluValValAsnSerGlySerGlu	417
Db	1083	GACCACGGCGCAACCCCTTACCCAGTAGAGCCATGAGTGGTCAACAGCGCTCTGAG	1142
OY	418	ValTrpTyrSerAsnGlyProGlyIleuLeuValIleAspCysAlaSerLeuGluIleGys	437
Db	1143	GTCGTGATCACCAATGGGGCCGGGGCTCTCTGTATGACACTGTCCCTCGGAGATGTGC	1202
OY	438	ArgArgLeuGluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlu	457
Db	1203	AGGGGGCTGGAGCCCTTACATGGCCCTTCATGTTACGTACGTGTGTGACGCTCTAG	1262
OY	458	GlyArgGlyGluGluValValTrpCysLeuAspAspLysAlaAsnSerLeuValMetCyr	477
Db	1263	GGCAGAGAGGGAGGGAGCTGTCTGGTGGCTGGAGAACAGGCCAACCTCTGTGTATGTAC	1322
OY	478	HisSerThrThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArg	497
Db	1323	CACCTCCACACCTTACACAGCTGTGTGCCCCGTACTTCTGGGGGGTCCCCAGCCCTTCAG	1382
OY	498	AspMetPheProValAlaArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsn	517
Db	1383	GACATGTTTCCCGTGGCGCCCTTGGACACGGAACCCCGGAGCCACACGGCCAC	1442
OY	518	ProLysValProGluGlyAspSerIleAlaAspValSerIleMetCysSerGluGluLeu	537
Db	1443	CCAAAGGGCTGTAGGGGGACTCATCTCGCGGACGTAGACATCATGTATGACGTAGGAGCTG	1502
OY	538	GlyThrGlnIleLeuIleHisGlnGluSerLeuThrAspTyrCysSerMetSerTyr	557
Db	1503	GGCAGCGAGATTCGATCCACCAAGGAATCCTACAGTACTAGTGCCTCATGTCTCTAC	1562
OY	558	SerSerSerProProAlaGlnAlaAlaValArgSerProSerSerLeuProSerSerProAla	577
Db	1563	TCCATCATCCCCACCCGCCAGGGCTGCAGAGTCCCTCCCTCAAGCCTCCCCAGCTCCCCACA	1622
OY	578	SerSerSerSerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrPro	597
Db	1623	AGTCTTCTCACTGTGCTTCTCTCCACCGACTGGGAGAGATTCACAGATGTTACATATGCC	1682
OY	598	GlyAlaAlaSerAspArgSerGluHisAspLeuThrProMetAspGlyGlyIuThrPheSer	617
Db	1683	GGTGCTCTCCCTCCAGAGTCTGAGCATGTACCTGACCCCATGAGCGGGGAGACCTTTCAGC	1742
OY	618	GlnHisLeuGlnAlaValLysIleLeuAlaValArgAspLeuIleThrValProArgArg	637
Db	1743	CACGACCTTGGCGCCGTGAAGAGTCTCGCGGTAGAGACTCATTTTGGGTCCCCAGGGGC	1802
OY	638	GlyGlyAspValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyValArgVal	657
Db	1803	GGTGAAGATGTATTCGTCATTTGGCTGCGAGAAAGATTTGAAGCCACAGCGGGCGGAGTTC	1862
OY	658	IleAlaValLeuLysAlaAlaArgGluLeuThrProHis-Gly-----	670
Db	1863	ATTTCGCTCTTAAAGCCCAAGAGCTGACTCCGCAATGGGATTTATGCGCACTCTCTGTGG	1922


```

Db 1082 AACTCTGTGTGATGTACCACTCCACCACTACCAAGCTGTGTGCCCCGCTACTTTCGCGG 1141
QY 492 VALProSerProLeuAlaGlyAspMetPheProValArgProLeuAspThrGluProProAla 511
Db 1142 GTCCCAAGCCCCCTCAGAGCATGTTTCCGTGGGCCCCCTTGAGACAGGAACCCCCGGGCA 1201
QY 512 AlaSerHisThrAlaAsnProLysValProGluGluAspSerIleAlaAspValSerIle 531
Db 1202 GCCAGCCACAGCGCCACCCCAAGAGTGGCTGAGGGGAGCTCATCGCCGAGGTGAGCATC 1261
QY 532 MetTyrSerGluGluLeuGluThrGlnIleLeuIleHisGlnGlnSerLeuThrAspTyr 551
Db 1262 ATGTACACTGAGAGCTGGGACGAGATCCTGATCCACAGAAATTCATCTAGTACTAC 1321
QY 552 CysSerMetSerTyrSerSerSerProProArgGlnAlaAlaArgSerProSerSer 571
Db 1322 TGTTCATGTCTCTCTACTCTCTCATCCGACCCGCGGAGGTGGCAAGTCCCTCCAGAC 1381
QY 572 LeuProSerSerProAlaSerSerSerSerValProPheSerThrAspCysGluAspSer 591
Db 1382 CTCGCCAGCTCCACAGCAAGTTCTTCCAGTGTGCTTCTCCACCCGAGTCCGAGAGCTCA 1441
QY 592 AspMetLeuHisThrProGluAlaAlaSerAspArgSerGlnHisAspLeuThrProMet 611
Db 1442 GACTGTCTACTACGCCCCGGGTGCTGCTCCGACAGGTGTGAGCATGAGCTGACCCCATG 1501
QY 612 AspGluGluThrPheSerGlnHisLeuGlnAlaValLysIleLeuAlaValAlaArgAspLeu 631
Db 1502 GACGGGAGACCTTCAGCAGACACCTGACGCGGTGAAGATCCTGCGCTCAGAGACCTC 1561
QY 632 IleThrValProAlaGluGlyGlyAspValIleValIleGluLeuGluLysAspSerGlu 651
Db 1562 ATTTGGGTCCCGACGCGGT -GGAGATGTTATGTCATGAGCTGAGAAAGATCTGAA 1620
QY 652 AlaGlnArgGlyArgValIleAlaValLeuLysAlaArgGluLeuThrProHisGlyVal 671
Db 1621 GCCACAGCGGCGGCGAGATATGCGCTTTAAAGCCCCAGAGCTGACTCCGCAATGGGTG 1680
QY 672 LeuValAspAlaAlaValAlaLysAspThrValValCysThrPheGluAsnGluAsn 691
Db 1681 CTGTGTGATGCTGCGGTGTGGCAAGACACTGTGTGTGCACCTTTGAATAATGAAAC 1740
QY 692 ThrGluTyrCysLeuAlaValAlaTyrArgGlyTyrProGluAlaArgGluPheAspIlePheTyr 711
Db 1741 ACAAGATGTCCTGCGGTGTGGAGGGCTGGGGCGCCAGGAGATTCGACATTTTCTAC 1800
QY 712 GlnSerTyrGluGluLeuGlyArgLeuGluLysLysThrArgLysArgArg 728
Db 1801 CAGTCCACAGAGAGAGCTGGCGGCTGAGAGCTTGCACTCCGCAAGAGAGG 1851

```

RESULT 3

US-10-144-771-18122
Sequence 18122, Application US/10144771

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig
TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
FILE REFERENCE: C1001321
CURRENT APPLICATION NUMBER: US/10/144, 771
CURRENT FILING DATE: 2002-05-15
NUMBER OF SEQ ID NOS: 47235
SEQ ID NO 18122
LENGTH: 1330
TYPE: DNA
ORGANISM: HUMAN
US-10-144-771-18122

Alignment Scores:

Pred. No.: 3,31e-50 Length: 1,330
Score: 785.00 Matches: 174
Percent Similarity: 51.70% Conservative: 8
Best Local Similarity: 49.43% Mismatches: 9
Query Match: 20.54% Indels: 162
DB: 9 Gaps: 2

```

US-09-836-392-21 (1-728) x US-10-144-771-18122 (1-1330)
QY 1 MetLeuArgHisLeuArgValAlaThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 20
Db 758 ATGCTGGGCGACCTGAGAGCCATGATGCCATGAAAGAACTTCTAGATTTCCGACAGAG 817
QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
Db 818 GCCAGACATGTCATGCTTGCACATCCCTGATCGTGTGATCAGTCCAGATCCGATCAGCATC 877
QY 41 HisProLeuCysPheAlaLeu -GluLeuAlaProLeuSerSerLeuAsnThrValLeuSe 60
Db 878 CACCGGCTGCTGCTGCGCTGGAGCTGGCCCCCTGGGCGAGTGTCAACACACAT - 931
QY 60 GlnAlaAsnAlaArgAspSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAl 80
Db 931 - 931
QY 80 ATGCTGlnAlaSerGlyLeuAlaTyrLeuHisLysAsnIleIlePheCysAspLe 100
Db 931 - 931
QY 100 ULysSerAspAsnIleLeuValTyrPheLeuAspValLysGlnHisIleAsnIleLysLe 120
Db 931 - 931
QY 120 UserAspTyrGlyLieserArgGlnSerPheHisGluGluValAlaLeuGluValGluGlu 140
Db 931 - 931
QY 140 rProGluTyrGlnAlaProGluIleArgProAlaGluValTyrAspGluLysValAspMe 160
Db 932 - 933
QY 160 rPheSerTyrGlyMetValLeuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlnHis 180
Db 934 GTTCTCTACGCGCATGGTGTATACAGCTGCTCAGAGAGCGTCCGTCGACGTGGGCGCA 993
QY 180 SHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnGlnPr 200
Db 994 CCACCACTTACATGTCGCAAGAAAGTTCTCCAAAGGCACTCCGCCAGTCTCGGGGCGAGCC 1053
QY 200 OGluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGlnCysTyrAspThrLysPr 220
Db 1054 GGAGGAAGTTCAGTTCATCGCTGACGAGCCCTCATGATGGAG - 1096
QY 220 OGluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
Db 1096 - 1096
QY 240 rPheMetTyrGluLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlnGln 260
Db 1096 - 1096
QY 260 uTyrThrValAlaPheThrAspGlyLysGluGlnSerArgAsnTyrThrValAlaAsnThr 280
Db 1097 - 1116
QY 280 rGluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
Db 1117 AGAGAAAGGCTTCTGGAGAGACAGAGATCACTGCGCGGAGTGAAGTGAAGCTGCA 1176
QY 300 nLeuGlnValGlnArgSerLeuTyrPheAlaThrGlnAspGlnLysIleTyrIleTyrTh 320
Db 1177 GCTCAAGTCCAGACGCTCAGCTGATAGCCAGGAGACACAGAAATCTATATCTACAG 1236
QY 320 rLeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVa 340
Db 1237 CCTTAAGGAGTATGTGCCATTAACGTCGCCACAGAGCCCTGAGACACCCGCGCTGTGCT 1296
QY 340 lThrCysPheLeuAlaValAlaProValIleLysLys 351

```

Db 1297 CACGTGTTCTTGCGAGTACTGTATCAAAAG 1330

RESULT 4

US-09-912-293-127829

Sequence 127829, Application US/09912293

GENERAL INFORMATION:

APPLICANT: Rosen, et. al.

TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100

FILE REFERENCE: PO-100

CURRENT APPLICATION NUMBER: US/09/912,293

CURRENT FILING DATE: 2001-07-26

PRIOR APPLICATION NUMBER: 08/103,744

PRIOR FILING DATE: 1993-08-09

PRIOR APPLICATION NUMBER: 09/249,651

PRIOR FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: 08/104,507

PRIOR FILING DATE: 1993-08-09

PRIOR APPLICATION NUMBER: 08/196,363

PRIOR FILING DATE: 1994-02-15

PRIOR APPLICATION NUMBER: 09/859,490

PRIOR FILING DATE: 2001-05-18

PRIOR APPLICATION NUMBER: 08/196,362

PRIOR FILING DATE: 1994-02-15

PRIOR APPLICATION NUMBER: 08/221,623

PRIOR FILING DATE: 1994-03-31

PRIOR APPLICATION NUMBER: 08/220,691

PRIOR FILING DATE: 1994-03-31

PRIOR APPLICATION NUMBER: 09/741,830

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 09/813,155

PRIOR FILING DATE: 2001-03-21

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 244538

SEQ ID NO 127829

LENGTH: 352

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (351)..(351)

OTHER INFORMATION: n is equal to a, t, g, or c

US-09-912-293-127829

Alignment Scores:

Pred. No.: 1.52e-36

Score: 596.00

Percent Similarity: 99.14%

Best Local Similarity: 99.14%

Query Match: 15.598

DB: 6

Length: 352

Matches: 115

Conservative: 0

Mismatches: 1

Indels: 0

Gaps: 0

US-09-836-392-21 (1-728) x US-09-912-293-127829 (1-352)

QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuIsthrProGlyAlaAla 600

Db 2 AGGTGCTCTTCCACTGCTGCGAGGAGTCAATGCTAATGCGCGGTGCTGCC 61

QY 601 SerAspArgSerGluIstAspLeuThrProMetAspGlyGluThrPheSerGlnIstLeu 620

Db 62 TCCGAGAGGCTGAGCATGACCTGACCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAT 121

QY 621 GlnAlaValIstLeuAlaValArgAspLeuIstIrrPValProArgArgIstGlyAsp 640

Db 122 CAGCGCTGTAAGATCTCGCGCTGCAAGAGCTCATTTGGTCCCGAGCGGTGAGAT 181

QY 641 ValIleValIleGlyLeuGluIstAspSerGluAlaGlnArgGlyArgValIleAlaVal 660

Db 182 GTTATGCTGATGCTGCGGAGGAGGATCTGCGCGCGAGGAGGAGGAGGAGGAGGAT 241

QY 661 LeuIstAlaArgGluLeuThrProIstGlyValIleValAspAlaAlaValAlaIst 680

Db 242 TTTAAAGCCGAGAGTGAAGTCCGCGATGGGCTGCTGTGTGATGCTGCGCGGTGCGCAAG 301

QY 681 AspThrValIstCysThrPheGluAsnGluAsnThrGluTrpCysLeu 696

Db 302 CACTGTTGTGTGCACTTTGAAATGAAACAGAGAGTGTGCTG 349

RESULT 5

US-10-349-781-16014

Sequence 16014, Application US/10349781

GENERAL INFORMATION:

APPLICANT: Selhammer, Jeffrey J.; Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.; Stuve, Laura L.

APPLICANT: Mullaly, Sara J.; Naughton, Rebecca E.

TITLE OF INVENTION: POLYNUCLEOTIDES OF CELLS AND TISSUES OF THE FEMALE REPRODUCTIVE

FILE REFERENCE: PD-1028-2 CON

CURRENT APPLICATION NUMBER: US/10/349,781

CURRENT FILING DATE: 2003-01-21

PRIOR APPLICATION NUMBER: 09/540,764

PRIOR FILING DATE: 2000-03-30

PRIOR APPLICATION NUMBER: 08/992,868

PRIOR FILING DATE: 1997-12-11

PRIOR APPLICATION NUMBER: 60/032,838

PRIOR FILING DATE: 1996-12-13

PRIOR APPLICATION NUMBER: 08/734,050

PRIOR FILING DATE: 1996-10-18

PRIOR APPLICATION NUMBER: 60/006,111

PRIOR FILING DATE: 1995-10-24

PRIOR APPLICATION NUMBER: 08/733,814

PRIOR FILING DATE: 1996-10-18

PRIOR APPLICATION NUMBER: 60/005,864

PRIOR FILING DATE: 1995-10-26

PRIOR APPLICATION NUMBER: 08/763,920

PRIOR FILING DATE: 1996-12-11

PRIOR APPLICATION NUMBER: 60/008,794

PRIOR FILING DATE: 1995-12-14

PRIOR APPLICATION NUMBER: 08/772,783

PRIOR FILING DATE: 1996-12-23

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 61458

SOFTWARE: PERL Program

SEQ ID NO 16014

LENGTH: 288

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No: hu00910748

NAME/KEY: unsure

LOCATION: 192, 263

OTHER INFORMATION: a, t, c, g, or other

US-10-349-781-16014

Alignment Scores:

Pred. No.: 1.47e-29

Score: 503.00

Percent Similarity: 97.92%

Best Local Similarity: 97.92%

Query Match: 13.16%

DB: 9

Length: 288

Matches: 94

Conservative: 0

Mismatches: 2

Indels: 0

Gaps: 0

US-09-836-392-21 (1-728) x US-10-349-781-16014 (1-288)

QY 394 IleAlaAspGluAspAlaArgGlnAsnProTyrProValIstAlaMetGluValIstAsn 413

Db 1 ATCCGCGATGAAGACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 60

QY 414 SerGlySerGluValIstIrrPThrSerAsnGlyProGlyIstLeuValIstAspCysAlaSer 433

Db 61 AGCGGCTCTGAGGCTGAGTACACAAATGGCGCGGCTCTCTTTCATCGAGTGTGCTCC 120

QY 434 LeuGlnIstCysArgArgLeuGluIstProTyrMetAlaProSerMetValThrSerValIst 453

Db 121 CTGAGATCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGATGAGTACGTCG 180

```

Oy      454 CyssterseargtugtiyaaggtgylugluvalValatPCysleuspaspAllysalaasner 473
Db      181 TGcAGCTGTGANGGCACAGGGAGGAGTGTGTGTGTGTCCTTGATGTGCAAGCCCAACTCC 240
Oy      474 LeuValMetIYrHISserThrThTYrGlnLeuCysAlaArgTyRpe 489
Db      241 TTGGTAGTGTACCACTCCACCACNACTACCACTGTGTGCCCGGTACTTC 288

RESULT 6
US-09-912-293-231260
; Sequence 231260, Application US/09912293
GENERAL INFORMATION:
APPLICANT : Rosen, et. al.
TITLE OF INVENTION : Human Genes, Sequences, and Expression Products 100
FILE REFERENCE : PO-100
CURRENT APPLICATION NUMBER : US/09/912,293
CURRENT FILING DATE : 2001-07-26 --
PRIORITY APPLICATION NUMBER : 08/103,744
PRIOR FILING DATE : 1993-08-09
PRIOR APPLICATION NUMBER : 09/249,651
PRIOR FILING DATE : 1999-02-12
PRIOR APPLICATION NUMBER : 08/104,507
PRIOR FILING DATE : 1993-08-09
PRIOR APPLICATION NUMBER : 08/196,363
PRIOR FILING DATE : 1994-02-15
PRIOR APPLICATION NUMBER : 09/659,490
PRIOR FILING DATE : 2001-05-18
PRIOR APPLICATION NUMBER : 08/196,362
PRIOR FILING DATE : 1994-02-15
PRIOR APPLICATION NUMBER : 08/221,623
PRIOR FILING DATE : 1994-03-31
PRIOR APPLICATION NUMBER : 08/220,691
PRIOR FILING DATE : 1994-03-31
PRIOR APPLICATION NUMBER : 09/741,830
PRIOR FILING DATE : 2000-12-22
PRIOR APPLICATION NUMBER : 09/813,155
PRIOR FILING DATE : 2001-03-21
Remaining prior application data removed - See file wrapper or PAM.
NUMBER OF SEQ ID NOS : 244538
SEQ ID NO 231260
LENGTH: 500
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY : misc_feature
LOCATION : (9)..(9)
OTHER INFORMATION : n is equal to a,t,g, or c
FEATURE:
NAME/KEY : misc_feature
LOCATION : (22)..(22)
OTHER INFORMATION : n is equal to a,t,g, or c
FEATURE:
NAME/KEY : misc_feature
LOCATION : (120)..(120)
OTHER INFORMATION : n is equal to a,t,g, or c
FEATURE:
NAME/KEY : misc_feature
LOCATION : (222)..(222)
OTHER INFORMATION : n is equal to a,t,g, or c
FEATURE:
NAME/KEY : misc_feature
LOCATION : (252)..(252)
OTHER INFORMATION : n is equal to a,t,g, or c
FEATURE:
NAME/KEY : misc_feature
LOCATION : (294)..(294)
OTHER INFORMATION : n is equal to a,t,g, or c
FEATURE:
NAME/KEY : misc_feature
LOCATION : (341)..(341)
OTHER INFORMATION : n is equal to a,t,g, or c
FEATURE:
NAME/KEY : misc_feature

```

```

1 LOCATION: (370)..(370)
2 OTHER INFORMATION: n is equal to a,t,g, or c
3
4 FEATURE:
5 NAME/KEY: misc_feature
6 LOCATION: (406)..(406)
7 OTHER INFORMATION: n is equal to a,t,g, or c
8
9 FEATURE:
10 NAME/KEY: misc_feature
11 LOCATION: (430)..(430)
12 OTHER INFORMATION: n is equal to a,t,g, or c
13
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: (435)..(435)
17 OTHER INFORMATION: n is equal to a,t,g, or c
18
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: (437)..(437)
22 OTHER INFORMATION: n is equal to a,t,g, or c
23
24 FEATURE:
25 NAME/KEY: misc_feature
26 LOCATION: (450)..(450)
27 OTHER INFORMATION: n is equal to a,t,g, or c
28
29 FEATURE:
30 NAME/KEY: misc_feature
31 LOCATION: (456)..(456)
32 OTHER INFORMATION: n is equal to a,t,g, or c
33
34 FEATURE:
35 NAME/KEY: misc_feature
36 LOCATION: (470)..(470)
37 OTHER INFORMATION: n is equal to a,t,g, or c
38
39 FEATURE:
40 NAME/KEY: misc_feature
41 LOCATION: (473)..(473)
42 OTHER INFORMATION: n is equal to a,t,g, or c
43
44 FEATURE:
45 NAME/KEY: misc_feature
46 LOCATION: (477)..(477)
47 OTHER INFORMATION: n is equal to a,t,g, or c
48
49 -S-09-912-293-231260

```

Alignment Scores:

Pred. No.:	1,04e-27	length:	500
Score:	483.00	Matches:	100
Percent Similarity:	90.43%	Conservative:	4
Best Local Similarity:	86.96%	Mismatches:	8
Query Match:	12.64%	Indels:	3
DB:	6	Gaps:	0

US-09-836-392-21 (1-728) x US-09-912-293-231260 (1-500)

QY 252 ALApheserSerGingIngInGInUrrThrValValAlpHeTrpaSpGlyLySgInGu 271U

Db 2 GCCTTCCTTCATCATCCAGGGGNCAGAGATACCCCTGGTGTGTTGGATGGAAAAAGAGAG 61

QY 272 SerTrgaSnTrThrValValAlaSnThGluSgLyLeuMetGluValGlnAArgMetCys 291U

Db 62 TCCAGGAACTACACAGGGGTGTAACACAGAGAAGGGCTCATAGGAGGTGACAGAGATGTC 122U

QY 292 CysProGlyMetLysValSerCysGlnLeuGlnValGlnArgSerLeuTrpThrAlaThr 311U

Db 122 TGCCCTGGAGTGAAGGTGAGTGTGCCCACTCCAGGTCCAGAGATCCCTGTGTGACAGCCACC 181U

QY 312 GluAspGlnLysIleTyrIleTyrThrLeuLysGlyMetCysProLeuAsnThrProGln 331U

Db 182 GAGGACCAAGAAATCTCATCTATACCTCAAGGCGATGTCCTTAAACACACCCCAA 241U

QY 332 GlnAlaLeuAspThrProAlaValValThrCysPhe-LeuAlaValProVal-IleLysL 351U

Db 242 CAGCGCTGGTGAATCTCCAGCGTGTGTACACGCTTCCTTGGCCGCTGTTNATTAATA 301U

QY 351 ysaSn-SerTyrLeuValLeuAlaGlyLeuAlaAspLys 363

Db 302 AGGATTCCTTCCTTCGCATCCCAAGGTTCATCAATGGG 340

RESULT 7

US-10-311-034-31

Sequence 31, Application US/10311034

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.
 APPLICANT: YUE, Henry
 APPLICANT: LAL, Preeti
 APPLICANT: BANDMAN, Olga
 APPLICANT: BOROWSKY, Mark L.
 APPLICANT: AU-YOUNG, Janice
 APPLICANT: LU, Yan
 APPLICANT: GANDHI, Ameena R.
 APPLICANT: TRIBOULEY, Catherine M.
 APPLICANT: CHAMLA, Narinder K.
 APPLICANT: YAO, Monique G.
 APPLICANT: LU, Dyung Alina M.
 APPLICANT: GREENWALD, Sara R.
 APPLICANT: RAMKUMAR, Jayalaxmi
 APPLICANT: GRIFFIN, Jennifer A.
 APPLICANT: KEARNEY, Liam
 APPLICANT: BURFORD, Neil
 APPLICANT: NGUYEN, Daniel B.
 APPLICANT: TANG, Y. Tom
 APPLICANT: BAUGHN, Mariah R.
 APPLICANT: HE, Ann
 APPLICANT: THORNTON, Michael
 APPLICANT: HAFALIA, April
 APPLICANT: ARVIZO, Chandra S.
 APPLICANT: GURURAJAN, Rajagopal
 APPLICANT: LO, Terence P.
 APPLICANT: KHAN, Farrah A.
 APPLICANT: RECIPON, Shirley A.
 APPLICANT: AZIMZAI, Yalda
 APPLICANT: POLICKY, Jennifer L.
 APPLICANT: DING, Li
 APPLICANT: GREYHER, Megan
 APPLICANT: ELIOTY, Vicki S.
 APPLICANT: THANGAVELU, Kavitha
 APPLICANT: BATRA, Sajeev
 APPLICANT: ISON, Craig H.
 TITLE OF INVENTION: HUMAN KINASES
 FILE REFERENCE: PI-0125 PCT
 CURRENT APPLICATION NUMBER: US/10/311,034
 CURRENT FILING DATE: 2002-12-10
 PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
 60/228,056
 PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-07-25
 NUMBER OF SEQ ID NOS: 52
 SOFTWARE: PERL Program
 SEQ ID NO 31
 LENGTH: 2693
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID No: 1242491CB1
 US-10-311-034-31

Alignment Scores:

Pred. No.: 4,45e-15 Length: 2693
 Score: 330.50 Matches: 108
 Percent Similarity: 47.348 Conservative: 79
 Best Local Similarity: 27.348 Mismatches: 150
 Query Match: 8.658 Indels: 59
 DB: 9 Gaps: 14

US-09-836-392-21 (1-728) x US-10-311-034-31 (1-2693)

QY 10 Alameltyasnbheserjgu-----Pheargnglualasermet 23
 DB 162 GCGTGAAGATTTTAAATACATACATCACTGAGGCTGTTAAGACAGAGCTTGCGTG 221

QY 24 LeuHsAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIleHisProLeu 43
 DB 222 CTTTGCCACCTTCACACCCCGCAGTTGATATCTTGCTGGCAGCTGGAGTTGCTCCCGG 281
 QY 44 CysPheAlaLeuGlnLeuAlaProLeuSerIleAsnThrValLeuSerIleAsnAla 63
 DB 282 ATGTTGGTGAATGAGTTCGCCCAAGGGTCTTGATGGCTGCTGCTGACAGAGCAAA 341
 QY 64 ArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAlaIleGlnIle 83
 DB 342 GCCAGC-----CTCACTAGAACCTCACACAGACAGAGATTCACCTCCACGTA 386
 QY 84 AlaSerGlyLeuAlaIleThrIleLysLysAsnIleIlePheCysAspLeuSerAsp 103
 DB 387 GCTGATGTTTGAATGATCTCCACGACCATGATTTATACGAGACCTGAAACCCAC 446
 QY 104 AsnIleLeuValIlePheSerLeuAspValLysGlnHisIleAsnIleLysLeuSerAsp 123
 DB 447 AATGTCGCTCTTTCACACTGATATCCCATGCTGGCATCATTTGCAAAAGATTGCTGAC 506
 QY 124 GlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnGlyThrProGlyTyr 143
 DB 507 GGCATTCCTCAGTACTGCTGTAAGATGGGATAAACATCAGAGGCGCACACAGGCTTT 566
 QY 144 GlnAlaProGlnIle---ArgProArgIleValIleAspGlnLysValAspMetPheSer 162
 DB 567 CGTGACCCGGAAGTTCGCCAGGAATGTCATTATTAACCAACGAGCTGATGTTATTTCA 626
 QY 163 TyrGlyMetValLeuTyrGlnLeuLeu---SerGlyLysArgProAlaLeuGly----- 179
 DB 627 TTGGTTTACTCTATATACATTTTGACACTGGAGGATGAAATAGTAGAGTTGAAG 686
 QY 180 -----HisIleGlnLeuGlnIleValLysLysLeuSerIleArgPro 195
 DB 687 TTTCCAATGAGTTGATGATTAAGAAATACAGAAATACCTGATCAGTTAAGAA 746
 QY 196 ValLeuGlyGlnProGlnIleValGlnPheArgArgLeuGlnAlaLeuMetGlnCys 215
 DB 747 TATGTTGTGCCCA-----TGCCCTATGGTGAAGAAATTAATTAACAGTGT 794
 QY 216 TrpAspThrLysProGlnLysArgProLeuAlaLeuSerValValSerGlnMetLysAsp 235
 DB 795 TTGAAGAAATTCCTCAAAAGAGCTTCTGCGCCAGCTCTTACATTTGAATTC 854
 QY 236 ProThrPheAlaThrPheMetTyrGlnLeuCysGlyLysGlnThrAlaPhe----- 253
 DB 855 GCT-----GAATTAGTCTGCT--GACGAGACGCAATTTATTATAC 892
 QY 254 -----PheSerSerGlnGlyGlnIle-----Tyr 261
 DB 893 TAAAAAGTAATTTGTAATGCATGTTGCTACACACACAGAGGATGCAACAT 952
 QY 262 ThrValValPheTrpAspGlyLysGlnGlnSerArg-AsnTyrThrValValAsnThrGln 281
 DB 953 TTGGCTGGGCTGTGGCACAACGACAGAGAGAGCTCTATTTCTTGACTTAATACTGA 1012
 QY 281 uLysGlyLeuMetGlnVal-----GlnArgMetCysCysProGlyMetLysValSer 298
 DB 1013 AGGATACACTTCGAGGAAGTTGCTGATAGTAATATGTGCTTGAACCTTG----- 1064
 QY 298 rCysGlnLeuGlnValGlnArgSerLeuTrp-----ThrAlaThrGlnAspLysIle 316
 DB 1065 -GTGCATCTTCCCTGTGAAGAAAGAGATGGATGTGTCTGGGACAGAGCTGTAATCT 1123
 QY 316 eTyrlIleTyrThrLeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThr 336
 DB 1124 CCTGGTCATCAATACCGAAGATGGGAAAGAGACATACCTTGAAAGATGCTGATTC 1183
 QY 336 rProAlaValValThrCysPheLeuAlaValProValIleLys-----LysAsnSerTy 354
 DB 1184 T-----GTCACTTGTGTGATTCGATTCCTTTTCCAAAGCAAAACAAACAAA 1234


```

! TYPE: DNA
! ORGANISM: homo sapiens
US-60-426-964-1

Alignment Scores:
Pred. No.: 1,69e-13
Score: 311.00
Percent Similarity: 35.448
Best Local Similarity: 23.188
Query Match: 8.148
DB: 11 Gaps: 29

US-09-836-392-21 (1-728) x US-60-426-964-1 (1-3138)

OY 19 GlnGlnAlaSerMetLeuHisAlaLeuGlnHisProCysAlaValAlaLeuIleGlyLe 38
Db 590 CAGGAAACCCCGGGCTTTGGAGCCCTGGACAGACCCCAACAAATGAGCCCTTAGGGGCGCC 649
OY 39 SerIleAspPro-----LeuGlyPheAlaLeuGlnAlaProLeuSerSerLeuAsn 56
Db 650 TGCCTCAACCCCAACACCTCTGCTCAATGATGAGATGATGATGCCGGGGTGGTGCACATGAC 709
OY 57 ThrValIleuSerGlnAlaAlaArgSerSerPheIleProLeuGlnHisMetLeuThr 76
Db 710 AGGGTGTGGCAGGTGCCCGC-----GTGCCA---CTGCAGTGTGTGTC 751
OY 77 GlnIlyIleAlaTyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLys----- 93
Db 752 ---ACATGGGCTGTGGACGTGGCCCGGGGCAATGAACTACCTACACAAAGATGCCCTGTG 808
OY 94 AsnIleIlePheCysAspIleLysSerAspAsnIleLeuValTyrSerLeuAspValLys 113
Db 809 CCCATATTCACACCGGAGACCTCAATGCCAATCAACATCTGATC-----CTGAGGCCATC 862
OY 114 GlnHisIleAsn-----IleLysLeuSerAspTyrGlyIleSerArgGln 128
Db 863 GAGAAACCAACACCTGGGAGACAGCGGTGCTCAAGATCAAGATCGAGATTTGGCTGGCCCGGAG 922
OY 139 SerPheHisGlnGlyAlaLeuGlyValGlnGlyThrProGlyTyrGlnAlaProGluIle 148
Db 923 TGGCACAAGACCAACCAAGATGACGCTGGCGGGGACCTTACCTCGATGATGGCCCGGAGATT 982
OY 149 ArgProAlaGlyLeuValTyrAspGlnLysValAspMetPheSerTyrGlyMetValLeuTyr 168
Db 983 ATCCGCTCTCCCTCTCTCCATCAAGCATGATGATGTGGAGCTTGGGGGTGTGCTGTGG 1042
OY 169 GluLeuLeuSerGlyGlnArgProAlaLeuGlnHisHisGlnLeuGlnIleAla----- 186
Db 1043 GAGCTGCTGACGGGGGAGAGTCCCTACCGAGATCGAGACGCTTGGCCGTGGCGATGAC 1102
OY 187 -----LysLysLeuSerLysGlyIleArgProValLeuGlnProGluGluVal 203
Db 1103 GTGGCTATGAAATAGCTGAGCGTGCACATTCCTCCACATGCGCCGAGACC----- 1153
OY 204 GlnPheArgArgLeuGlnAlaLeuMetMetGlnCysTyrAspThrLysProGluLysArg 223
Db 1154 --TTTGCCGCC-----CTCCGTGAGGATGTGTGGACCCAGACCCCAAGCGGGCG 1201
OY 224 ProLeuAlaLeuSerValLysGlnMetLysAspProThrPheAlaThrPheMetTyr 243
Db 1202 -----CCAGATTTGGTGGACATCTTGAAG 1225
OY 244 GluLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGluTyrThrVal 263
Db 1226 CGGCTGAAAGCATCAAGCAACAGTCAACGCTTCCCAATGACCACTGAGTCTTCACATCG 1285
OY 264 ValPheTyrAspGlyLysGlnGluSerArgAsnTyr-----ThrValValAsnThrGlu 281
Db 1286 CTCGAGGAAGACTGGAAGCTGAGATTCAGACACATGTTGATGACATCTTGGACCAAGAG 1345
OY 282 LysGlyLeuMetGluValGlnArgMetCysCys-----ProGlyMet 295
Db 1346 AAGGAGCTTCGAGCCCTGAGAGAGAGACTCTGCGGGCGACAGAGACAGCGCTTCCAG 1405

```

Oy	295	tlYvalSerCys	-----glnleuGlnAlaGlnArgSerleuTPTPr			309
Db	1406	GAGAGCAGCTGGCGGCGCGGAGCAGAGCTGGCAGAACGTGAATGGACATCGTGGAA				1465
Oy	309	-----	-----	-----	-----	309
Db	1466	CGGAGCTGCACCTGCATGTGCACGTAGCCAGACGAGAACGCCCGGGGTCCGCAAGCC	-----	-----	-----	1525
Oy	310	-----AlaThrGlnAspGlnIleTyrIleTyr	-----ThrLeuIysGlnIleCys			325
Db	1526	AAGGGCACTTCAACGACGCGTTCTTCACGTAACGGAAGGACGACGCCACATCAGCTG				1585
Oy	325	spIleuAsnThrProGlnIleAlaLeuAspThrProAlaValAlaThrCysPheLeuAl				345
Db	1586	CCCTCTGGCTTTGACCTAATGATCACAAGTCCAGCGCT	-----CTCCA	1627	-----	1627
Oy	345	avalProvalIleIleIysAsnSerTyrIleuValIleuAlaGlyLeuAlaAspGlyLeuVa				365
Db	1628	ACTGTGGATAG-CGGAAAGATCCGATGGGGCCAGCCCGCTGCAGCCCGACATCAT				1666
Oy	365	lAlaValPheProValValArgIleTyrProIysAspSerCysSerTyrLeuCysSerHl				385
Db	1687	CCCCGGCTAGGGCCATTGTGCTGACTCCCGGAGC	-----TGGGGG	1731	-----	1731
Oy	385	spHAlaAsnArgSerIysPheSerIleAlaAspGlnAspAlaArgGlnAspTyrPr				405
Db	1732	CAGCAGAGTGGCAGCAGCAGTGGAGGAAGTGGACATGTAGCCCGGTGGGCC--CC	1788	-----	-----	1788
Oy	405	ovalIys	-----	-----	-----	407
Db	1789	AAAGAAAGAAAGACTGTGCGGGGCAGAAAGAAAGACAGACAGTGGGGCCAGCTCCAC	1848	-----	-----	1848
Oy	408	-----	-----AlaMetGlnValAlaAsnSerGlySerG			417
Db	1849	CCGTGACAGAGAGCGGGTGGAGAGAGAGAGAGCTGAAGGGCTGGGGAGGAAGCAAA	1908	-----	-----	1908
Oy	417	uValTPTPTySerAsnGlyProGlyLeu	-----LeuValIl	429	-----	429
Db	1909	ACAATGGATCAATCAAGTCCCCCAACTGGGCAAGTCCCCCAACACACACCATCAGCCCC	1968	-----	-----	1968
Oy	429	aspCysAlaSerIleuGlnIleCysArgArgLeu	-----	-----	-----	440
Db	1969	TGGCTTGGCCAGCTCATATAGATGGAGAGGTCCGGGAGGACAGAGATGGAGCGACAG	2028	-----	-----	2028
Oy	441	-----GlnProTyrMetAlaProSerMetValThrSerValValCysSerSerG	457	-----	-----	457
Db	2029	CGTCCCGCTTCCCCCTATCTGCAGACCCCGCTCACTCACTAGGCCACATCCCTCGGACCC	2088	-----	-----	2088
Oy	457	uGlyArgGlyGlnIleuValTyrCysLeuAspArgPylAlaAsnSerLeuValMetTy	477	-----	-----	477
Db	2089	CTCCCCGGGGCGCGGGCGCGGTGGAGCGGACGCGTCCGCGCCCCCGCTCGGTGGGG	2148	-----	-----	2148
Oy	477	hIleSerThrThrTyrGlnLeu	-----CysAla	486	-----	486
Db	2149	ACAGCGGCGCGGGGCGCGTGCAGACGTGGCGCTGATAGGTGGCGGCACGTCTGGGGC	2208	-----	-----	2208
Oy	487	-----	-----Ar	487	-----	487
Db	2209	TGTGGCGCTGGCGCCGACGTGGCGGAGCGCGCGCGCGCGGACGTGAGAGACAGCGCG	2268	-----	-----	2268
Oy	487	gTyPheCysGlyValProSerProLeuArgAspMetPheProValArgProLeuAspTh	507	-----	-----	507
Db	2269	CTGGCTGCAGCGGCTTCTTTTCCCGCGCGCGCGCTCCCG--CGGGGCTC	2320	-----	-----	2320
Oy	507	glnuProProAlaAlaSerHsThrAlaAsnProIysValProGlnGlyIysAspSerIleAl	527	-----	-----	527
Db	2321	-AGCCACCCGCGGTCCCGACGAGCGCGCGGAAAGACGTGGGGCGCGCTGGGCTGGC	2379	-----	-----	2379
Oy	527	aspValSerIleMetIysIserGlnIleuGlyThrGlnIleuIleuGlnIleuGlnIleu	547	-----	-----	547
Db	2380	GCCC-----TGGCGACCCCTGTGTCTGCTGTCTGCTC	2409	-----	-----	2409


```

Db 1737 GGGGAGAGAACAAAGAGTGTATCAAGTCCCAACTGGGAGATGCCCAACAA 1796
Oy 427 -----LeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu----- 440
Db 1797 CACACCATCGCCCTGGCTTCCGCCAGCTCAATGATGAGAGATTGCGGAGGAGCA 1856
Oy 441 -----GluProTyrMetAlaProSerMetValThrSerVa 452
Db 1857 GGATGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1916
Oy 452 IValCysSerSerGluArgIleGluGluValValTyrCysLeuAspAspAlaAs 472
Db 1917 ACTGCTGCGGAGCCCTCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1976
Oy 472 nSerLeuValMetTyrHisSerThrTyrGlnLeu-----CysAl 486
Db 1977 CCCCCCTGCTGGGAGACAGCGCCCGGGGGGGGGGGGGGGGGGGGGGGGGGG 2036
Oy 486 a----- 486
Db 2037 CACGCTGTGGGGGCTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2096
Oy 487 -----ArgTyrPheCysGlyValProSerProLeuArgAspMetPheProVa 502
Db 2097 TGAGGAGAGCAGCGCGCTGGCTGCGAGCGCTCTTCTTCCCGGGGGGGGGGGGG 2154
Oy 502 IArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysValProG1 522
Db 2155 -CGGGGGCTC-----AGCCCAACCGCGCGCGCGCGCGCGCGCGCGAGAGCTGGGGCC 2207
Oy 522 uGluAspSerIleAlaAspValSerIleMetTyrSerGluGluGluValTyrGlnIle 542
Db 2208 CGGCGCTGGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2237
Oy 542 uIleHisGlnIleSerLeuThrAspTyrCysSerMetSerTyrTyrSerSerProPr 562
Db 2238 CGTGTGCGTGTGCTGCGTGTGCTGCGTGTGCTGCGTGTGCTGCGTGTGCTGCG 2297
Oy 562 oArgGlnAlaAla---ArgSerProSerSerLeuProSerSerProAlaSerSerSe 581
Db 2298 TGACAGAGCGCGACAGCGCGCGCGCTCCACACCGCGCGCGCGCGCGCGCGAC 2357
Oy 581 rValProPheSerThr-----AspCysGluAspSerAspMetLeuHisThrPr 597
Db 2358 GCCCTCGCGCCAGCACACACCGCGCTGGTGAAGCTGAGAGAGCTTCAAGAGAGCC 2417
Oy 597 oGly-----AlaAlaSerAspArgSerGluH1 606
Db 2418 CGGCGAGTGTGCTGAGCGCCAGCCAGCTCAAGCTGTATGCGGTGTGAGCGCGGG 2474
Oy 606 sAspLeuThrProMetAspGly 613
Db 2475 CCGCGGAGCGCATCGATGGG 2496

```

```

RESULT 12
PCT-US02-24567-86
Sequence 86, Application PC/TUS0224567
GENERAL INFORMATION:
APPLICANT: Chinnaiyan, Arul M.
APPLICANT: Rubin, Mark A.
APPLICANT: Sreekumar, Arun
TITLE OF INVENTION: Expression Profile of Prostate Cancer
FILE REFERENCE: OM-07221
CURRENT APPLICATION NUMBER: PCT/US02/24567
CURRENT FILING DATE: 2002-12-23
PRIOR APPLICATION NUMBER: US 60/309,581
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/334,468
PRIOR FILING DATE: 2001-11-15
NUMBER OF SEQ ID NOS: 123
SOFTWARE: PatentIn version 3.2
SEQ ID NO 86
LENGTH: 3435

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-24567-86

Alignment Scores:
Pred. No.: 2,44e-12 Length: 3435
Score: 296.50 Matches: 174
Percent Similarity: 35.03% Conservative: 88
Best Local Similarity: 23.26% Mismatches: 276
Query Match: 7.76% Indels: 211
Gaps: 29

US-09-836-392-21 (1-728) x PCT-US02-24567-86 (1-3435)
Oy 19 GlnGluAlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIle 38
Db 718 CAGGAAGCGCGGCTCTTGGAGCCCTGGAGCCCGCCAGCCCAACATATATGCTTGGGGGGCGCC 777
Oy 39 SerIleHisPro-----LeuGlyPheAlaLeuGluLeuAlaProLeuSerSerLeuAsn 56
Db 778 TGCCCTAACCCCCCAGACCTGTGCTAGTATGAGATATGCCCCGGGGTGGTGCACCTGAGC 837
Oy 57 ThrValLeuSerGluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThr 76
Db 838 AGGAGTGTGGCAGAGTGGCGCG-----GTGCCA---CTTCAGCTGTGGTGC 879
Oy 77 GlnLysIleAlaTyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLys----- 93
Db 880 ---AACTGGGCTGTGACAGTGGCGCGCGGCGCAGAACCTACACAAATGATGCCCTGTG 936
Oy 94 AsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValTyrSerLeuAspValLys 113
Db 937 CCGATATCCACCGGAGACCTCAAGTCCATCAACATCTGTATC-----CTGAGGGCATGC 990
Oy 114 GlnHisIleAsn-----IleLysLeuSerAspTyrGlyIleSerArgGln 128
Db 991 GAGAACACAACTCGCGACAGACAGCGGTGCTCAAGATCAGCAGACTTGGCGCTCCCGCGAG 1050
Oy 129 SerPheHisGlnGlyAlaLeuGlyValGluGlyThrProGlyTyrGlnAlaProGluIle 148
Db 1051 TGGCACAAGACCAACCAAGATGAGCGCTGGCGGAGACTTACCGCTGATGCGCGGAGGTT 1110
Oy 149 ArgProHisGlyIleValTyrAspLysValAspMetPheSerTyrGlyMetValLeuThr 168
Db 1111 ATCCGTCTCTCCCTCTCTCCAAAAGCAGTATGCTGTGAGCTGGGGTGGCTGCTGCTGG 1170
Oy 169 GluLeuLeuSerGlyGlnArgProAlaLeuGlyHisGlnLeuGlnIleAla----- 186
Db 1171 GAGCTGTGACGGGGAGAGTCCCTTACCGCTAGATGAGACCGCTTGGCGCTGGCGTATGGC 1230
Oy 187 -----LysAlaLeuSerLysGlyIleArgProValLeuGlyGlnProGluGluVal 203
Db 1231 GTGGCTATGATTAAGCTGACGCTGCCATTCCTCCACGCTGCCCGAGGCC----- 1281
Oy 204 GlnPheArgArgLeuGlnAlaLeuMetGluCysTyrAspThrLysProGluLysArg 223
Db 1282 ---TTTGCCGCG-----CTCCTGAGGATGTGTGGAGCCAGACCCAGCGGGCGG 1329
Oy 224 ProLeuAlaLeuSerValAlaSerGlnMetLysAspProThrPheAlaThrPheMetTyr 243
Db 1330 -----CCAGATTTCGCTGACATCTTGAAG 1353
Oy 244 GluLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGluTyrThrVal 263
Db 1354 CGGCTTGAAGTATGATGAGACAGTACAGCCGCTTCCAGATGCCACTGAGAGCTTCCACTCG 1413
Oy 264 ValPheThrAspGlyLysLeuGluLysArgAsnTyr-----ThrValValAsnThrGlu 281
Db 1414 CTGCAGAGAGACTGGAGAGTGAAGATTCAGCAAGTGTGATGACCTTCGAGCAAGAGAG 1473
Oy 282 LysGlyLeuMetGluValGlnArgMetCysCys-----ProGlyMet 295
Db 1474 AAGGAGCTTTCGAGAGCGCTGAGAGAGAGCTGTCCGGGGCGGACAGAGAGAGCGCTTCCAG 1533

```


PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 1327

LENGTH: 3143

TYPE: DNA

ORGANISM: Human

US-09-949-016-1327

Alignment Scores:

Pred. No.: 3,08e-12 Length: 3143

Score: 294.50 Matches: 174

Percent Similarity: 35.03% Conservative: 88

Best Local Similarity: 23.26% Mismatches: 276

Query Match: 7.71% Indels: 211

DB: 7 Gaps: 29

US-09-836-392-21 (1-728) x US-09-949-016-1327 (1-3143)

19 GlnGlnAlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyLe 38
 591 CAGAGAGCCCGGCTCTTGAGCCCTGCAGCACCACATATATGCTTGGGGCGCC 650
 39 SerIleHisPro-----LeuGlnHisAlaLeuGlnHisProLeuSerSerLeuAsn 56
 651 TGCCTCAACCCCGCCACACCTGCTGCTAGTGTGAGTGTGCGGGGTGTGCTGAGC 710
 57 ThrValLeuSerGlnHisAlaArgPheSerPheIleProLeuGlnHisMetLeuThr 76
 711 AGGCTGCTGCAGTCCGCGCGG-----GTGCCA---CCTCAGCTGTGCTG 752
 77 GlnHisAlaLeuGlnHisAlaSerGlyLeuAlaLeuHisGlyLeu 93
 753 ---AAGTGGCTGTGAGTGTGCGCGGCGCATGACATACATACATACATGATGATGATG 809
 94 AsnIleIlePheCysAspLeuHisSerAspAsnIleLeuValIlePheSerLeuAspValys 113
 810 CCCATCATCAACCGGACCTCAATGATCATCAATCATCTGATC-----CTGAGGCGCATC 863
 114 GlnHisIleAsn-----IleLeuSerAspPheGlyIleSerArgGln 128
 864 GAGAACCAACACCTCGCAGACAGCTGCTCAAGATCAGCAGCTGCGCTGCGCGCGAG 923
 129 SerPheHisGlnGlyAlaLeuGlnGlyValGlnGlyThrProGlyThrGlnAlaProGlyLeu 148
 924 TGGCACAAGACACCAAGATGAGCGCTGCGGGGACCTAGCGCTGATGCGCGGAGGT 983
 149 ArgProAlaGlnIleValIleAspGlnHisValAspMetPheSerIleGlyMetValLeuThr 168
 984 ATCCGCTCTCCCTCTCTCCAAAGCAGTGTGTGAGCTTGTGCGGGGTGCTGCTG 1043
 169 GlnLeuLeuSerGlnGlnArgProAlaLeuGlnHisGlnLeuGlnIleAla----- 186
 1044 GAGCTGCTGAGCGGGGAGGTCCCTACCGTGAAGATGAGCGCTTGGCGGCGATGAGC 1103
 187 -----LysLeuSerLeuGlnGlyLeuArgProValLeuGlnGlnProGlnVal 203
 1104 GTGGTATGATATAGCTACGCTGCTTCTCCACGCTGCGCGCGAGCC----- 1154
 204 GlnPheArgGlnGlnAlaLeuMetGlnCysIlePhePheThrIleProGlnValArg 223
 1155 ---TTGCGCGC-----CTCTGAGAGATCTGTGAGCCAGACCCCGAGCGGG 1202
 224 ProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThrPheMetThr 243
 1203 -----CCAGATTTCGAGCATCTTGAAG 1226
 244 GlnLeuCysGlyGlyGlnThrAlaPhePheSerSerGlnGlnGlnIleThrVal 263
 1227 CGGCTTGAATGATCAAGACGACCGCTGTTCAGATGACCACTGAGATCTCTCACTG 1286
 264 ValPheThrAspGlyGlyGlnGlnLeuSerArgAsnThr-----ThrValValAsnThrGln 281

1287 CTGCAGAAAGACTGGAACTGGAGATTCAGACATGTTTGTATGACTTGGACCAAGAG 1346
 282 LysGlyLeuMetGln-ValGlnArgMetCysCys-----ProGlyMet 295
 1347 AAGAGCTTGGAGCGCTGTGAGAGGAGCTGTGCGGGCGGCGACAGACAGCGCTTCAG 1406
 295 LysValSerCys-----GlnLeuGlnValGlnArgSerLeuThrAlaThrGln 312
 1407 GAGGAGCAGCTGCTGCGCGCGGAGAGAGAGTGGCAGAACTGATGATGAGCA----- 1458
 312 uAspGlnIleIleThrIleThrIleLeuGlnGlyMetCysProLeuAsnThrProGlnI 332
 1459 -----TCGTGAGAGCGGAGCTGACCTGCTCATGTGCTGCGAGCTG 1496
 332 nAlaLeuAspThrPro----- 337
 1497 AGCCAGAGAGAACCCCGGCTCCGACAGCGCAAGGCAACTTCAGACCGCGCTGCTC 1556
 338 -----AlaValValThrCysPheLeuAlaValProValIleLeuGly 351
 1557 AAGTGGCGGAGAGCGCGGAGCAGCATACCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 1616
 351 sAsnSerIleuValLeuAlaGlyLeuAlaAspGlyLeuValAlaValPhePro----- 369
 1617 GTCCAGGCTCTCCAACTGTGATAGCGGAGAAAGATCCATGAGGCGCGCGCGCTGCTG 1675
 370 -----ValValArgGlyThrProLysAspSerCysSe 380
 1676 AAGCCCAAGCATATCCCGCGCTGAGGCGCATTCGCTGCTGCTGCTGCTGCTGCTG 1727
 380 rTyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlnAlaAla 400
 1728 -----TGTGTGAGCAGACAGATGAGCAGAGCAGAGTGTGAGGAGATGAGATGAGAGCG 1780
 400 gGlnAsnProIleThrProValLys----- 407
 1781 CGGTGGGCGC---CCAAAGAGAGAAAGAACTGTGTGCGGGGCAAGAGAGAGCAACCTG 1837
 408 -----AlaMetGlnValVal 412
 1838 GGGGCGCGCTCCACCTCGCAGAAAGAGCGGCTGGAGAGAGAGAGAGAGAGAGAGAGCT 1897
 412 IAsnSerGlySerGlnValIlePheSerAsnGlyProGlyLeu----- 426
 1898 GGGGAG 1957
 427 -----LeuValIleAspCysAlaSerLeuGlnIleCysArgArgLeu----- 440
 1958 CACAACCGAGTGGCGCTGCTGCTGCGCAGCTCATGATGAGAGAGAGAGAGAGAGAG 2017
 441 -----GlnProIleMetAlaProSerMetValIleThrSerV 452
 2018 AGATGAG 2077
 452 alValCysSerSerGlnGlnArgGlnGlnValIleThrCysLeuAspAspValAla 472
 2078 CACTGCTGCGGAGCGCTCCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2137
 472 snSerIleValMetIleHisSerThrThrIleGlnLeu-----CysA 486
 2138 CCGCGCGTGGTGGGAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2197
 486 la----- 486
 2198 CCAGCTGCTGCGGCGCTGCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 2257
 487 -----ArgThrPheCysGlyValProSerProLeuArgAspMetPheProV 502
 2258 GTAG 2316
 502 aArgProLeuAspThrGlnProProAlaAlaSerHisThrAlaAsnProLysValProG 522

```

Db 2317 --CGGGGCTC-----AGCCACCGCGCGCTCCACGGCCCGCCGGAAGCTGGCC 2368
QY 522 LuGIyAspSerIleAlaAspValSerIleMetYrSerGluLeuGlyThrGlnIleL 542
Db 2369 CCGGCGCTGGCGCTGGCGCC-----TCGGCCACCC 2398
QY 542 euIleHsGlnGluSerLeuThrAspTyrCysSerMetSerTyrSerSerProp 562
Db 2399 TCGTGTCCGTGCTGCTCGCTCCGACTGCAACTCACAGCGCTCATCTGCTGCTGACA 2458
QY 562 roHrGlnAlaAlaAspSerProSerSerLeuProSerProAlaSerSerSererv 582
Db 2459 GTACAGAGCGCCAGCGCGCGCCCTCCACACCCCTCCCGCGCCGCCACACCCA 2518
QY 582 alPro-----PheSerThrAspC 588
Db 2519 CGGCGCTCGCCAGACACCAACCCCTGTGTGAGCTGAGAGTTCAGAGAGAC 2578
QY 588 ysGluAspSer---AspMetLeuHsThrProGlyAlaAlaSerAspArgSerGluHsA 607
Db 2579 CCGCGCAGTGTGCTACAGCCCGCCAGCTGACGCTGCTGCTGAGCGCGGACCC 2638
QY 607 spleuHrProMetAspGly 613
Db 2639 GCGGAGCGCATCGATGGG 2658

RESULT 15
US-60-455-1697
Sequence 1697, Application US/60455444
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: BEGOVICH, Ann
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001455
CURRENT APPLICATION NUMBER: US/60/455,444
NUMBER OF SEQ ID NOS: 50986
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1697
LENGTH: 3143
TYPE: DNA
ORGANISM: Homo sapiens
US-60-455-444-1697

Alignment Scores:
Pred. No.: 3 08e-12 Length: 3143
Score: 294.50 Matches: 174
Percent Similarity: 35.03% Conservative: 88
Best Local Similarity: 23.26% Mismatches: 276
Query Match: 7.71% Indels: 211
DB: 10 Gaps: 29

US-09-836-392-21 (1-728) x US-60-455-444-1697 (1-3143)
QY 19 GlnGluAlaSerMetLeuHsAlaLeuGlnHsProCysIleValAlaLeuIleGlyLe 38
Db 591 CAGGAACCCCGGCTTGTGAGCCCTCAGACCCCAACATAATGCGCTTAGGGGCGCC 650
QY 39 SerIleHsPro-----LeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsn 56
Db 651 TGCCTCAACCCCGCCAGACCTGCTGCTAGTATGAGATGACCGCGGCTGTGACAGC 710
QY 57 ThrValLeuSerGluAlaAlaArgAspSerSerPheIleProLeuGlyHsMetLeuThr 76
Db 711 AGGCTGTGCGAGCTGCCCG-----GTGCCA---CCTCAGTGTGCTGTC 752
QY 77 GlnLysIleAlaTyrGlnIleAlaSerGlyLeuAlaTyrLeuHsIleLysLys----- 93
Db 753 ---AAGCGGCTGTGCGAGCGCGCGGCGCATGACTACACATGATGCCCTGTG 809
QY 94 AsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValTyrSerLeuAspValLys 113

```

```

Db 810 CCATATATCCACCGGAGCTCAAGTCCATACATCTGATC-----CTGAGGCCATC 863
QY 114 GlnHsIleAsn-----IleLysLeuSerAspTyrGlyIleSerArgGln 128
Db 864 GAGAACCAACCTCCGACAGACGGTGTCAATATCAGGACTTGGCCCTCGCCGCGAG 923
QY 129 SerPheHsGlnGluAlaLeuGlyValGluGlyThrProGlyTyrGlnAlaProGluIle 148
Db 924 TGGCACAAGACACCAAGATGAGCGCTGCGGGAGCTACACCTGATGAGCGCGAGGTT 983
QY 149 ArgProAlaGlyValTyrAspGluLysValAspMetPheSerTyrGlyMetValLeuTyr 168
Db 984 ATCCGCTCTCCCTCTTCCAAAGCAAGATGTGTGGAGCTTGGGCTGTGTGG 1043
QY 169 GluLeuSerGlyGlnArgProAlaLeuGlyHsIleGlnLeuGlnIleAla----- 186
Db 1044 GAGCTGTGAGGGGGAGAGTCCCTACCGGATGACACCTTGGCGTGTGGCTATGGC 1103
QY 187 -----LysLysLeuSerLysGlyIleArgProValLeuGlnProGluGluVal 203
Db 1104 GTGGCTATGATTAAGTACGAGCTGACCATTCCTCCAGTGTGCCCGAGCC----- 1154
QY 204 GlnPheArgArgLeuGlnAlaLeuMetLeuGlySerTrpAspThrLysProGluLysArg 223
Db 1155 ---TTTGGCCGC-----CTCCTGAGAGAAATGTGGAGCCAGACCCCGCGG 1202
QY 224 ProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThrPheMetTyr 243
Db 1203 -----CCAATTTGGTACACTTGAAG 1226
QY 244 GlnLeuCysCysGlyGlnThrAlaPheSerSerGlnGlyGlnIleTyrThrVal 263
Db 1227 CGGCTGAATCATCAAGACAGTACGAGCTTCCAGATGACAGTGCCTTCCACTCG 1286
QY 264 ValPheTrpAspGlyLysGluGluSerArgAsnTyr-----ThrValAlaSerThrGlu 281
Db 1287 CTCGAGAAAGCTGGAAGCTGAGATTCAGCAATGTTGATGACCTTCGACCAAGAG 1346
QY 282 LysGlyLeuMetGlu-ValGlnArgMetCysCys-----ProGlyMet 295
Db 1347 AAGGAGCTTGGAGCGCTGAGAGAGAGCTGCTCGGCGRACACAGAGACGCTTCCAG 1406
QY 295 LysValSerCys-----GlnLeuGlnValGlnArgSerLeuThrAlaThrGln 312
Db 1407 GAGGAGAGCTGCGCGCGGAGACAGAGCTGACAGAAAGTGAAGAGACA----- 1458
QY 312 uAspGlnLysIleTyrIleTyrThrLeuLysGlyMetCysProLeuAsnThrProGlnI 332
Db 1459 -----TCGTGGAAGGAGCTGACCTGCTCATGTGCTGACACTG 1496
QY 332 nAlaLeuAspThrPro----- 337
Db 1497 AGCCAGAGAAGCCCGGCTCCGCAAGCGCAAGGCAACTTCAAGCCAGCCGCTGCTC 1556
QY 338 -----AlaValValThrCysPheLeuAlaValProValIleLysLys 351
Db 1557 AAGCTGCGGAAGCGCGACACATGACCTGCGCTGTGAGTAAAGATACAA 1616
QY 351 sAsnSerTyrLeuValLeuAlaGlyLeuAlaAspGlyLeuValAlaValPhePro----- 369
Db 1617 GTCCAGCGCTCCCAACTGTG-GATAAGCGGAAGATCCGATGGGCGCAAGCCCTGTC 1675
QY 370 -----ValValArgGlyThrProLysAspSerCysSe 380
Db 1676 AAGCCAGACATATCCCGGCTGAGAGCGCATTCGCTGCTGCTGAGC----- 1727
QY 380 rTyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaAr 400
Db 1728 -----TGTGGGCGAGCAGAGAGTGCACAGAGAGTGAAGAGAGTGAAGTGAAGCCG 1780
QY 400 gGlnAsnProTyrProValLys----- 407
Db 1781 CGGTGGGCGCC---CNAAGAGAGAACTGTGCGGGGCGAGAGAAAGGAGAGAACGTG 1837

```

```
OY 408 -----AlaMetGluValVa 412
Db 1838 GGGGCCAGCTCCACCTCGAAGAGAGCGGGTGGAGAGAGAGAGAGAGAGCTGAAGGGGCT 1897
OY 412 IAsnSerGlySerGluValTyrPyrSerAsnGlyProGlyLeu----- 426
Db 1898 GGGGGAAGAGCAAGACAGTGTATCATAGTCCCCCAACTGGGCAAGTCCCCCAACA 1957
OY 427 -----LeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu----- 440
Db 1958 CACACCCAGTGGCCCTGGCTTGGCCACCTCATGAGATGAGAGAGTTCGGAGAGCAG 2017
OY 441 -----GluProTyrMetAlaProSerMetValThrSerV 452
Db 2018 AGGATGAGAGCAGCAGCGTGGCCCTTCCCTACTGACCCCGTCTACTGCTCAGTGC 2077
OY 452 AlValCysSerSerGluGlyArgGlyGluValValTyrCysLeuAspAspIysAla 472
Db 2078 CACTGCTGCGGACCCCTCCCGGGGGCGCGGCGCGCTGGAGCCGAGCCGCTCCGCGC 2137
OY 472 snSerLeuValMetTyrHisSerThrThrTyrGlnLeu-----CysA 486
Db 2138 CCCCCGCTGGTGGGAGACAGCGCCCGCGCGCTGCGACCTGCGCTGAGGCTGCG 2197
OY 486 Ia----- 486
Db 2198 CCACGCTGCTGGGGGCTGTGGGCTGGGGCGGAGCGTGGCGGAGCGCGCGCGCGAGC 2257
OY 487 -----ArgTyrPheCysGlyValProSerProLeuArgAspMetPheProV 502
Db 2258 GTGAGAGACAGCGCGCTGCTGACGCGCTCTTCTTCCCGCGCGCGCTTCCCG - 2316
OY 502 aLArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProIysValProG 522
Db 2317 -CGGGGCTC-----AGCCACCCGCGCGCTCCCGACGCGCGCGCGAAGAGTGGGCC 2368
OY 522 IuGlyAspSerIleAlaAspValSerIleMetTyrSerGluGluLeuGlyThrGlnIleL 542
Db 2369 CCGGCTGGGCGCTGGCGCC-----TCGGCCACCC 2398
OY 542 euIleHisGlnSerLeuThrAspTyrCysSerMetSerTyrSerSerSerProp 562
Db 2399 TCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 2458
OY 562 roArgGlnAlaAlaArgSerProSerSerLeuProSerSerProIaSerSerSerServ 582
Db 2459 GTGACGAGGCGCAGCGCGCGCGCTCCACACACCTCCCGCGCGCGCCACACCCA 2518
OY 582 aLPro-----PheSerThrAspC 588
Db 2519 CGCCCTGCGCCACAGACACACCCCTGGTGGACCTGGAGACTGGAGAGCTTCAAGAAAGAC 2578
OY 588 ysgIuAspSer--AspMetLeuHisThrProGlyAlaAlaSerAspArgSerGluHisA 607
Db 2579 CCGCCAGTCTGCTGACGCGCACCGACGTCACGGCTGCTGAGCCGCGGCGAC 2638
OY 607 spleuThrPrometaspGly 613
Db 2639 GCGGAGCGCATCGGAYGG 2658
```

Search completed: April 15, 2003, 09:53:18
Job time : 562 secs

QY 169 ACCGTGCTCGAGAACGCGAGATTTCTTATACCCCTGGACACATCTCCACC 228
 |||||
 Db 184 ArgValLeuAlaGlyArg-----ValPro-----ProHisValLeuVal 197
 QY 229 CAAATAATAGCTTACAGATCGCTGGCGCTGACCTACCTGACACAGAAA----- 279
 |||||
 Db 198 ---AsnTrpAlaValGlnValAlaArgGlyMetAsnTrpLeuHisAsnAlaProVal 216
 QY 280 AACATCATCTTCTGTGACCTGAAATCGGACAAACATCTGGGTGTGTCCCTTGACGTACG 339
 |||||
 Db 217 ProIleHisArgAspLeuLysSerIleAsnIleLeuIle-----LeuGlnAlaIle 234
 QY 340 GAGCAATCAAC-----ATCAAGCTATCTGACTACGGATTTCCAGGACG 384
 |||||
 Db 235 GlnAsnHisAsnLeuAlaAspThrValLeuLysIleHisAspPheGlyLeuAlaArgGln 254
 QY 385 TCATTCATGAGGCGCCCTAGGCGGTGAGGCGACCTCTGCTACAGGCGCCAGATC 444
 |||||
 Db 255 TrpHisLysThrThrLysMetSerAlaIleGlyThrTrpAlaTrpMetAlaProGlnVal 274
 QY 445 AGGCTCGCATTTATATGATGAGAGATATATGTTCTCTCTATGAAATGTCTAC 504
 |||||
 Db 275 IleArgLeuSerLeuPheSerLysSerSerAspValTrpSerPheGlyValLeuLeuTrp 294
 QY 505 GAGTGTCTGTCAGACAGCGCCCTGACCTGGCGCACACAGCTCCAGATTGCC----- 558
 |||||
 Db 295 GluLeuLeuThrGlyValProTrpArgGluIleAspAlaLeuAlaValAlaIleTrpGly 314
 QY 559 -----AAGAAGCTCTCCAAAGGCGATCCGCGGCTTGGGGACCGGAGAGATG 609
 |||||
 Db 315 ValAlaMetAlaLysLeuThrLeuProIleProSerThrCysProGlnPro----- 331
 QY 610 CAGTTCGCGGACACTGAGCGGCTCATGATGAGTGGCTGGGCACTAACCAGAGAGCGA 669
 |||||
 Db 332 ---PheAlaArg-----LeuLeuGlnLysCysTrpAspProAspPheHisGlyArg 347
 QY 670 CCG-----CTGGCCCTGTCTGGGTGAGCGGAC 696
 |||||
 Db 348 ProAspPheGlySerIleLeuLysArgLeuGlnValIleGlnGlnSerAlaLeuPheGln 367
 QY 697 ATGAAGAGCCGACCTTTGGCACCTC----- 723
 |||||
 Db 368 MetProLeuGlnSerPheHisSerLeuGlnLysPTrpLysLeuGlnIleGlnHisMet 387
 QY 723 ----- 723
 Db 388 PheAspAspLeuArgThrLysGlnLysLeuArgSerArgGlnGlnLeuLeuArg 407
 QY 723 ----- 723
 Db 408 AlaAlaGlnGlnLysArgPheGlnGlnLysLeuArgArgArgGlnGlnLeuAla 427
 QY 724 -----ATGATGAACTGTGCTGGGAGAGACACA 753
 |||||
 Db 428 GlnArgGlnMetAspIleValGlnArgGlnLeuHisLeuMetCys-----GlnLeu 445
 QY 754 GCGTCTCTCATCCAGGCGGACAGAGACACCGGTGTTTGGAGAGAAAGAGAG 813
 |||||
 Db 446 SerGlnGlnLysProArgValArgLys-----AsnPheLysArgAlaVal 463
 QY 814 TCCAGGAACTACACGGTGT-----GAAAC----- 839
 |||||
 Db 463 AlLeuLysLeuArgGlnLysSerSerHisIleSerLeuProSerGlyPheGlnHisLysI 483
 QY 840 -----ACAGAAGCGCTCATGAGAGTGCAGAGG 867
 |||||
 Db 483 LeuHisValGlnAlaSerProThrLeuAspLysArgLysGlySerAspLysAlaSerProp 503
 QY 868 ATGCGCTGCC----- 878
 |||||
 Db 503 roAlaSerProSerIleIleProArgLeuArgAlaIleArgLeuThrProValaAspCysG 523

QY 879 -----TGGAGTGAAGTGAAGTGC 897
 |||||
 Db 523 LysGlySerSerSerGlySerSerSerGlyLysSerGlyThrTrpSerArgGlyProp 543
 QY 898 CAGCTCCA-----GCTCCAGAGATCC 918
 |||||
 Db 543 roLysLysGlnGlnLeuValGlyGlyLysLysGlyArgThrTrpGlyProSerSer 563
 QY 919 CTGAGACAGACACCGA-----GGACAGAAAATCTACATCTCAACCCCAAGGCGATG 972
 |||||
 Db 563 hreLysGlnLysGlnArgValGlyGlyLysGlnLysLeuLysGlyLysGlyLysSerL 583
 QY 973 TGCCCTT-----AAACACCCCAACAGGCC 999
 |||||
 Db 583 LysGlnTrpSerSerSerAlaProAsnLeuGlyLysSerProLysHisThrProIleAla 603
 QY 1000 TTGGATCTCCAGCTGTCGTCACCTGCTTGGCGCTGCTTATTAAGAATCC 1059
 |||||
 Db 603 roGlyPheAlaSer-Leu-----AsnGln 610
 QY 1060 TACCTGCTTACGCGGCGCTCGCATGGCTGTGGCTGTGCTTCCGTGTCGCGGCG 1119
 |||||
 Db 611 MetGlnLysPheAlaGlnAlaGlnAspGlyLysSerValProProSerProTrpSer 630
 QY 1120 ACCCAAGAGACAGCTGCTCTACCTGCTGCTACACACAGCAACAGCTCAAGTTACG 1179
 |||||
 Db 631 ThrPro-----SerTrpLeu-SerValProLeuProAlaGlnProSerProGln 646
 QY 1180 ATCCGAGTACAGACGAGCGAGACCCCTACCCAGTGAAGCGATGAGGTGTAC 1239
 |||||
 Db 646 ValAlaArgAlaProTrpLysu---ProThrProSerAla---ProProAlaArgTrpGlyHis 664
 QY 1240 AGCGCTCTGAG-----TGTGTACAGCAATGGCGCGGCG 1275
 |||||
 Db 664 sGlyAlaArgArgArgCysAspLeuAlaLeuLeuGlyCysAlaThrLeuLeuGlyAlaVala 684
 QY 1276 CTCCTGTGATGATGATGCTGCTCCCTGAGATGTGACAGCGCG-----TG 1320
 |||||
 Db 684 LysLysGlnLysAlaAspAlaIleGlnAlaArgAlaAspGlyGlnGlnLysArgArgTr 704
 QY 1321 G-----AGCCCTACATG 1332
 |||||
 Db 704 PLeuAspGlyLeuPhePheProAlaGlyArgPheProArgGlyLeuSerProProAla 724
 QY 1333 GCCCCTCATGATTACGTACG----- 1354
 |||||
 Db 724 aArgProHisLysArgArgGlnAspValGlyProGlyLeuGlyLeuAlaProSerAlaThr 744
 QY 1355 -----TCGTGT-----GCAGCTTGAGGCGAGA 1377
 |||||
 Db 744 LLeuValSerLeuSerSerValSerAspCysAsnSerThrArgSerLeuLeuArgSerAs 764
 QY 1378 GGGAGAGAGCTGCTGTGCTGCTGATGACAGGCCAATCTGTGTGATGATACATCC 1437
 |||||
 Db 764 PserAspGlnAlaIleAlaProAla-----AlaProSerPr 775
 QY 1438 ACCAGCTACAGCTGTGAGCGGTACTTGTGCGGCGCCAGCCCGGAGGAGCATG 1497
 |||||
 Db 775 OProProSerProProAlaPro-----ThrProThrProSerProSerThr 790
 QY 1498 TTTCCCG-----TCGGCGCTTGAGACAGAAACCCGCGAGCCAGCCAC 1542
 |||||
 Db 790 rasnProLeuValAspLeuGlnLysSerPheLysLysAspProArgGlnSerLeu-- 809
 QY 1543 ACGGCCAACCAAGGTGCTGAGAGGAGCTCCATCCGCGAGCTGACATCATATACGT 1602
 |||||
 Db 810 -----ThrProThrHisValThrAlaIleAlaCysAlaVala 820
 QY 1603 GAGAGCTGGGACAGACATCTGATCCACAGAAATCAGTCACTACTGCTCATG 1662
 |||||
 Db 820 LserArgLysHisArgArg-----ThrProSerAspLysAlaLeuGlyGlnArgGln 837
 QY 1663 TCTCTACTCTCATCCACCCCGCGAGGCTGCGAGGTCCCTCCAGGCTCCAGC 1722

```

Db 837 yProgluProAlaGlyHisGly-----ProGlyProArgAspLeuAerph 854
QY 1723 TCACCACAGATTCTTCACAGTGCCTTTCACACGACGCGGAGACTGACATGCTA 1782
Db 854 eProArgLeuProAspProGlnAlaLeuPheProAlaArgArgPro-----870
QY 1783 CACACGCGCGGTGCTGCTCCGACAGCTGAGCATGACCTGACCCCATGAGAGCGGAG 1842
Db 871 -----ProGluPheProGlyArg 876
QY 1843 ACCTTCAGCAGCAGCTGACGCGCGTGAAGATCTCGCCGTACAGACACTCATTTGGATC 1902
Db 876 gProThrThrLeuThrPheAlaPro-----884
QY 1903 CCCAGGCGCGGTGAGATGTTATGTCATTTGGCTGAGAGAAGATTCGAAGCCGACGG 1962
Db 885 -----ArgProArgPr 888
QY 1963 GCGCGAGTCAATTCGCTCTTAAAGCCGAGAGCTGACTCCGATGGGCTGCTGGAT 2022
Db 888 oAlaAlaSerArgProArgLeuAspPro-----Tply 899
QY 2023 GCTGCGCGGTGCGAAGACACTGTGTGTGACCT 2059
Db 899 sleuValSerPheGlyArgThrLeuThrIleSerPro 911

```

RESULT 10
US-10-369-022-56
Sequence 56, Application US/10369022

```

GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Rosenfeld, Julie Beth
TITLE OR INVENTION: METHODS AND COMPOSITIONS IN TREATING
TITLE OR INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
TITLE OR INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
TITLE OR INVENTION: 27410, 33760, 619, 15985, 69112, 2158, 224, 615, 44373,
TITLE OR INVENTION: 95431, 22345, 2387, 16658, 55054, 16314, 1675, 9569 OR
TITLE OR INVENTION: 13424 MOLENUMIN
FILE REFERENCE: MP102-0271RMONIM
CURRENT APPLICATION NUMBER: US/10/369, 022
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: US 60/360,495
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/370,121
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/373,010
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/373,908
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/377,717
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: US 60/379,949
PRIOR FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: US 60/382,409
PRIOR FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US 60/385,280
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 60/386,879
PRIOR FILING DATE: 2002-06-06
PRIOR APPLICATION NUMBER: US 60/387,536
PRIOR FILING DATE: 2002-06-10
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 56
LENGTH: 953
TYPE: PRT
ORGANISM: Homo sapiens
US-10-369-022-56

```

Alignment Scores:

```

Pred. No.: 1.6e-07 Length: 953
Score: 317.50 Matches: 200
Best Similarity: 31.968 Conservative: 86
Best Local Similarity: 22.358 Mismatches: 258
Query Match: 7.828 Indels: 351
DB: 6 Gaps: 38

```

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-369-022-56 (1-953)

```

QY 55 CAGAGGCGCAGCATGCTGCGACGCGCTGACACCCCGCATCGGCGGCATCGGCGATC 114
Db 144 GlnGluAlaArgLeuPheGlyAlaLeuGlnHisProGlnIleAlaLeuArgGlyAla 163
QY 115 AGCATCCACCG-----CTCTGCTGCGCCGTGAGCTGCGCGCGCTGACAGCGCTGAC 168
Db 164 CysLeuAsnProPheHisLeuValMetCylurAlaArgGlyAlaLeuSer 183
QY 169 ACCGTGCTGTCGAGAGAGCGGACAGATTCCTTTATACCCCTGGAGACATGCTCAC 228
Db 184 ArgValLeuAlaGlyArgArg-----ValPro-----ProHisValLeuVal 197
QY 229 CAAATAATAGCGTACAGATGCGCTGCGCGCTGCGCTGACGCGACAGAA-----279
Db 198 ---AsnTrpAlaValGlnValAlaArgGlyMetAsnTrpLeuHisAsnAspAlaProVal 216
QY 280 AACATCATCTTCTGTCGACCTGAGAGTGCAGACATCTGCTGCTGCTGCTGACGTCAG 339
Db 217 ProIleHisArgAspLeuPheLeuSerIleAsnIleLeu-----LeuGluAlaIle 234
QY 340 GAGCAGCATCAAC-----ATCAGATATCTGACATGAGGAGATTGCGAGCAG 384
Db 235 GlnAsnHisAsnLeuAlaAspThrValLeuValSerIleThrAspPheGlyLeuAlaArgGlu 254
QY 385 TCATTCATGATGAGGCGCGCGCTGAGGCGGACCTGCTGCTGCTGCTGCGCGCGAGATC 444
Db 255 TrpHisLeuThrThrLeuMetSerAlaAlaGlyThrAlaTrpMetAlaProGluVal 274
QY 445 AGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 504
Db 275 IleArgLeuSerLeuPheSerLeuSerSerAspValTrpSerPheGlyValLeuLeuTrp 294
QY 505 GAGTGTGCTGACAGACAGCGCGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 558
Db 295 GlnLeuLeuThrGlyGlyValProTyrArgGluIleAspAlaLeuAlaValAlaTyrGly 314
QY 559 -----AAGAGCTGTCCAGAGCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
Db 315 ValAlaMetAsnLeuLeuThrLeuProIleProSerThrCysProGluPro-----331
QY 610 CAGTTCGCGCGCTGACGAGCGCGCTGATGATGATGATGATGATGATGATGATGATGAT 669
Db 332 ---PheAlaArg-----LeuLeuGlnGlnLucyStrAspProAspProHisGlyArg 347
QY 670 CCG-----CGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 696
Db 348 ProAspPheGlySerIleLeuValArgLeuGlnValIleGlnGlnSerAlaLeuPheGln 367
QY 697 ATGAGAGCGCGCGATTTTGCACCTTC-----723
Db 368 MetProLeuGlnSerPheHisSerLeuGlnIleAspTrpIleGlnIleGlnHisMet 387
QY 723 -----723
Db 388 PheAspAspLeuArgThrLeuGlyGlyLeuArgSerArgGlnGlnLeuLeuArg 407
QY 723 -----723
Db 408 AlaAlaGlnGlnArgPheGlnGlnGlnLeuArgArgArgGlnGlnLeuAla 427
QY 724 -----ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
Db 428 GluArgGluMetAspIleValGlnArgGlnLeuHisLeuLeuMetCys-----GlnLeu 445

```


Db 144 GlnGluAlaArgPheUheGlnAlaLeuGlnHisProAsnIleIleAlaLeuArgGlyAla 163
 QY 115 AGCATCCACCGG-----CTGTGCTTCGGCCCTGAGCTGGCGCGCTCAAGACCTCAAC 168
 Db 164 CysLeuAsnProGlnHisLeuGlnValMetGlnTyrAlaArgGlyGlyAlaLeuSer 183
 QY 169 ACCGTGCTGCCAGAACCGCAGAGATCTTCCTTATACCCCTGGAGACAGTCTCAC 228
 Db 184 ArgValLeuAlaGlnArgArg-----ValPro---ProHisValLeuVal 197
 QY 229 CAAAAATAGCTTACACAGATCGCTGGCGCTGCTTACTGACACAGAAA----- 279
 Db 198 ---AsnTyrAlaValAlaGlnValAlaArgGlyMetAsnTyrLeuHisAsnAspAlaProVal 216
 QY 280 AACCATGCTCTGTGACCTGAGAGTGCAGACATCTGTGTGGTCCCTGACGTCAAG 339
 Db 217 ProIleIleHisArgAspLeuSerSerIleAsnIleLeuIle-----LeuGlnAlaIle 234
 QY 340 GAACACATCAAC-----ATCAGCTATCTGACTACAGGATTTGAGAGCAG 384
 Db 235 GlnAsnHisAsnLeuAlaAspThrValLeuLysIleThrAspPheGlnLeuAlaArgGln 254
 QY 385 TCATTCATGAGGGCGCCCGCTAGAGCGTGGAGGACATCCCTGACAGCGCCAGAGATC 444
 Db 255 TyrHisLysThrThrLysMetSerAlaAlaGlyThrTyrAlaThrPheAlaProGlnVal 274
 QY 445 AGGCTCGCATGTATATGATGAGAAGTAGATATGTCCTATGAGATGGTCTTAC 504
 Db 275 IleArgLeuSerLeuPheSerLysSerSerAspValTyrPheGlnLysIleLeuTyr 294
 QY 505 GAATGCTGTCTGAGACAGCGCTGCTGACCTGAGCGCCAGCAGCTCAGATTGCC----- 558
 Db 295 GlnLeuLeuThrGlnGlnValProTyrArgGlnIleAspAlaLeuAlaValAlaTyrGly 314
 QY 559 -----AAGAGCTGTCTCAAGCGCATCCGCCGCTTGGGGCAGCGCGAGAGAGT 609
 Db 315 ValAlaMetAsnLysLeuThrLeuProIleProSerThrCysProGlnPro----- 331
 QY 610 CAGTTCGGCGACTGACAGCGCTCATGATGATGATGATGATGATGATGATGATGATGAT 669
 Db 332 ---PheAlaArg-----LeuLeuGlnGlnCysTyrThrAspProAsnProHisGlyArg 347
 QY 670 CGC-----CTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
 Db 348 ProAspPheGlnSerIleLeuLysArgLeuGlnValIleGlnGlnSerAlaLeuPheGln 367
 QY 697 ATGAGAGACCCGACTTTGGACCTTC----- 723
 Db 368 MetProLeuGlnSerPheHisSerLeuGlnGlnAspTyrLysLeuGlnIleGlnHisMet 387
 QY 723 ----- 723
 Db 388 PheAspAspLeuArgThrLysGlnLysGlnLeuArgSerArgGlnGlnLeuLeuArg 407
 QY 723 ----- 723
 Db 408 AlaAlaGlnGlnGlnArgPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 427
 QY 724 -----ATGATGACTGTGCTGTGGAGACACACA 753
 Db 428 GlnArgGlnMetAspIleValGlnArgGlnLeuHisLeuMetCys-----GlnLeu 445
 QY 754 GCGCTTCTTCATCCAGGCGCAGAGTACACCGTGTGTTGGATGGAAGAGAGAG 813
 Db 446 SerGlnGlnLysProArgValArgLys-ArgLysGly-----AsnPheLysArgAlaVal 463
 QY 814 TCCAGACTACACAGTGTGT-----GACAC----- 839
 Db 463 AluLeuLysLeuArgGlnGlnLysSerSerHisLysSerLeuProSerGlyPheGlnHisLysI 483
 QY 840 -----AAGAGAGCGCTCATGAGAGCTCAGAGG 867
 Db 483 IethrValGlnAlaSerProThrLeuAspLysArgLysGlySerAspLysAlaSerProp 503

QY 868 ATGTGCTGCC----- 878
 Db 503 roAlaSerProSerIleIleProArgLeuArgAlaIleArgLeuThrProValAspCysG 523
 QY 879 -----TCGATGAAGTGAAGTGC 897
 Db 523 LysLysSerSerSerGlySerSerSerGlySerGlyThrThrPheSerArgGlyProp 543
 QY 898 CAGCTCCA-----GGTCCAGAGATCC 918
 Db 543 roLysLysGlnGlnLeuValGlyGlyLysLysLysGlyArgThrThrPheLysProSerSer 563
 QY 919 CTGTGACACCCACCA-----GACACGAAAATCTACATCTACACCTCAGAGGCATG 972
 Db 563 hrLeuGlnLysGlnArgValGlyGlyGlnGlnArgLeuLysGlyLeuGlyGlySerL 583
 QY 973 TGCCCTT-----AACACACCCCAACAGCC 999
 Db 583 LysGlnTyrSerSerSerAlaProAsnLeuGlyLysSerProLysHisThrProIleAla 603
 QY 1000 TTGATATCTCAGCTGTGCTGACCTGCTTGTGGCGCTGTATTAAAGAAATTC 1059
 Db 603 roGlyPheAlaSer-Leu-----AsnGln 610
 QY 1060 TACCTGCTTTACGGGCGCTCGCCGATGGCTGTGGCTGTTCCTCCGTGCGGGCG 1119
 Db 611 MetGlnGlnPheAlaGlnAlaGlnAspGlyLysSerSerValProProSerProTyrSer 630
 QY 1120 ACCCCAAAGACAGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179
 Db 631 ThrPro-----SerTyrLeu-SerValProLeuProAlaGlnProIleProGln 646
 QY 1180 ATCCGAGTGAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1239
 Db 646 YAlaArgAlaProTyrGln---ProThrProSerAla---ProProAlaArgTyrGlnH 664
 QY 1240 AGCGGCTCTGAGG-----TCTGTACACCAATGGCGCGCG 1275
 Db 664 sGlyAlaArgArgArgCysAspLeuAlaLeuGlnCysAlaThrLeuLeuGlnAlaVal 684
 QY 1276 CTCTTGTCTATCGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Db 684 LysLeuGlnLysAlaAspValAlaGlnAlaArgAlaAlaAspLysGlnGlnArgArg 704
 QY 1321 G-----AGCCTACATG 1332
 Db 704 PLeuAspGlyLeuPhePheProArgAlaGlyArgPheProArgGlyLeuSerProProAl 724
 QY 1333 GCCCCTCCATGTTGCTACGTACG----- 1354
 Db 724 aAcgProHisGlyArgArgGlnAspValGlyProGlyLeuGlyLeuAlaProSerAlaTh 744
 QY 1355 -----TCGTGT-----GCAGCTGTGAGGCGAGA 1377
 Db 744 rLeuValSerLeuSerSerValSerAspCysAsnSerThrArgSerLeuLeuArgSerAs 764
 QY 1378 GGGAGAGAGGTGCTGTGCTGCTGATGACAGAGCCAACTGCTGATGATGATGATGATGATG 1437
 Db 764 pSerAspGlnAlaAlaProAla-----AlaProSer 775
 QY 1438 ACACATTCACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1497
 Db 775 oProProSerProProAlaPro-----ThrProThrProSerProSerTh 790
 QY 1498 TTTCCTG-----TGCGCCCTTGAGACACGAAACCCCGGACAGCCAGCCAC 1542
 Db 790 rAsnProLeuValAspLeuGlnLeuGlnLysSerPheLysLysAspProArgGlnSerLeu 809
 QY 1543 ACGGCAACCAAGAGTGTGCTGAGGGGAGTCCATCCGCGAGCGTACATCATCATCATCATCAT 1602
 Db 810 -----ThrProThrHisValThrAlaLysAlaVal 820


```

OY 1603 GAGAGCTGGGCGAGATCTGATCCAGCAAGATCACTGACTGCTCCATG 1662
DB 820 lserarglyhiisargly-----ThrProserAspGlyAlaLeuLysGlnArgG1 837
OY 1663 TCCTCTACTCTCCATCCCGCCCGCCAGGCTCCAGGTCCTCCAGCTCCAGC 1722
DB 837 yProProGluProAlaGlyHisGly-----ProGlyProArgAspLeuAspPh 854
OY 1723 TCCCGAGCAAGTCTCTCCAGTGTGCTTCTCCAGCAGCTGCGAGCTCAGACATCTA 1782
DB 854 eProArgLeuProAspProGlnAlaLeuPheProAlaArgArgPro----- 870
OY 1783 CATACGCCCGCTGCTGCTCCGAGCTGAGCATGACCTGACCCCATGAGCGGAG 1842
DB 871 -----ProGluPheProGlyArg 876
OY 1843 ACCTTACCCAGCAGCTCCAGCGCCGAGAGATCCTGCGCCCTCAGACCTCATTTGGGTC 1902
DB 876 gProThrThrLeuThrPheAlaPro----- 884
OY 1903 CCCAGCGCGGTGAGATGTATGTCATTTGGCTGAGAAAGATTTGTAAGCCAGCGG 1962
DB 885 -----ArgProArgPr 888
OY 1963 GGCCGAGTATTCGCTCTTAAAGCCGAGAGCTGACTCGCATGGGCTGCTGGTAT 2022
DB 888 oAlaAlaSerArgProArgLeuAspPro-----Tply 899
OY 2023 GCTGCGGTGGCAAGGACACTGTTGTGCACT 2059
DB 899 sLeuValSerPheGlyArgThrLeuThrIleSerPro 911

RESULT 12
US-10-288-798-12
; Sequence 12: Application US/10288798
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Daniel B.
; APPLICANT: WALIA, Nalinder K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Marian R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dzung Alina M.; LAU, Preeti G.;
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program

```

```

; SEQ ID NO: 12
; LENGTH: 1097
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 55052990CD1
US-10-288-798-12

Alignment Scores:
Pred: No. 1,97e-07 Length: 1097
Score: 315.50 Matches: 198
Percent Similarity: 33.76% Conservative: 94
Best Local Similarity: 22.89% Mismatches: 275
Query Match: 7.77% Indels: 299
DB: Gaps: 35

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-288-798-12 (1-1097)
OY 7 AGGCACTGGCGGCGCAACCGATGCGATGAGAACTTCCGAGTCCGCGAGAGCGCAGC 66
DB 174 ArgHisAspProAspGluAspIleSerGlnThrIleGluAsnValArgGlnGluAlaLys 193
OY 67 ATGCTCAGCGCGCTGAGCAGCAACCCCTGCATCTGTCGCGCTCATCGCATCAGATCCACCG 126
DB 194 LeuPheAlaMetLeuLysHisProAsnIleIleAlaLeuArgGlyAlaCysLeuLysGlu 213
OY 127 -----CTCTGCTTCCGCCCTCGAGCTCGCGCGCTCAGACGCTCAACACCTGCTGTC 180
DB 214 ProAsnLeuCysLeuValMetGluPheAlaArgGlyProLeuAsnArgValLeuSer 233
OY 181 GAGAACGCCAGAGATTTCTTATACCCCTGCGACACATCTCTACCCCAAAATA--- 237
DB 234 -----GlyAsnArgIleProProAspIleLeu 242
OY 238 -----GCTTACAGATCGCTCGGCGCTGCGCTACCTGCACAGAAA----- 279
DB 243 ValAsnThrPalaValGlnIleAlaArgGlyMetAsnThrLeuLeuAspGluAlaIleVal 282
OY 280 AACATCATCTTCTGTGACCGGAGTGGGACAAATCTGTGTGGTCCCT----- 330
DB 263 ProIleIleHisArgAspLeuLysSerAsnIleLeuLeuLysValGluAsn 282
OY 331 ---GAGCTCAAGAGACATCAATCAATCAATCTGACTGACGAGATTTGAGGAGCTCA 387
DB 283 GlyAspLeuSerAsnLysIle-----LeuLysIleThrAspPheIleuAlaArgGluTrp 301
OY 388 TTCATGAGCGCGCTAGCGCTGAGGCGCATCTGCTTACCGAGCGCCGAGATCAGG 447
DB 302 HisArgThrThrLysMetSerAlaAlaGlyThrTyAlaArgPheAlaProGluValIle 321
OY 448 CCGCGCTTATATATATGTAAGAGTAGATATTTCTCCATGAGATGCTGCTACGAG 507
DB 322 ArgAlaSerMetPheSerLysGlySerAspValTrpSerTrpIleValLeuLeuTrpGlu 341
OY 508 TTCCTGTCAGCAGCGCCCTGCACTGCGCCAGCACACAGCTCAGATTC----- 558
DB 342 LeuLeuThrGlyGluValProPheArgGlyIleAspGlyLeuAlaValAlaTrpGlyVal 361
OY 559 -----AAGAAGCTGTCCAAAGGCGATCCCGCGTCTGGGCGAGCGGAGAGTGCAG 612
DB 362 AlaMetAsnLysLeuAlaLeuProIleProSerThrCysProGluPro----- 377
OY 613 TTCGGGGAATGAGGCGCTCATGATGAGATGCTGGACACTAAGCCAGAGAGCGG 672
DB 378 -----PheAlaLysLeuMetGluAspCysTrpAsnProAspProHisSerArg-- 393
OY 673 CTGAGCCCTGCTGGTGGAGCAGATGAAGAGCCGCTTTTCCACCTCATGTATGA 732
DB 394 -----ProSerPheThrAsnIleLeuAspGln 402
OY 733 CTGTGCTGTGGAGAGCAGACAGCGCTTCTTCTC-ATCCAGCGCCAGAGATACCGCTG 791
DB 733 -----

```



```

Db 668 .....GlnHisSerProSerGlnSerTyrLeuGlySerProPheProArgGly 683
OY 1302 .....GGAGATTCGACGCGCGCTGAGCCCTCAT-.....GGCCCCCCTCATGGT 1346
Db 684 GlnAspGlyAspGlyProSerSerAspGlyLeuHisGlnGluProThrProValAsnSer 703
OY 1347 TAGCTGAGTCGTGACGAGCTGAGAGG.....CAGAGGGAGAGAGTCTGCTG 1394
Db 704 AlaThrSerThrProGlnLeuThrProThrAsnSerLeuThrArgGlyGlyAla----- 721
OY 1395 GTGCTGTGATGACAGAGCCAACTCTGTGTGATGACCACTGACCACTGACAGCTGTG 1454
Db 722 .....HisHis-----ArgArg 725
OY 1455 TGCCCGGTACTTCTGCGGGGT.....-CCCCAGCCCGCT----- 1487
Db 726 CysGlnValAlaLeuLeuGlnGlyCysGlyAlaValLeuAlaThrGlyLeuGlyPheAsp 745
OY 1476 .....-CCCCAGCCCGCT----- 1487
Db 746 LeuLeuGlnAlaLeuGlyCysGlnLeuLeuProLeuGlnGluProGluProProAlaArg 765
OY 1488 .....CAGGACATGTTT-----CCCGTCCGCGCC---TTGACAC 1520
Db 766 GlnGluLysArgArgGlnGly-LeuPheGlnArgSerSerArgProArgArgSerThr 785
OY 1521 GGAGCCCCCGGACCGACG-----CACAGCGCCCAACCCAAAGTGCCTGAGGGGAGCTC 1574
Db 785 rSerProSerArgLysLeuPheLysLysGlnGluProMetLeuLeuLysLysPrr 805
OY 1575 CAGCGGAGAGTGCATGATGATGAGTGCAGAGGAGCTGGGACCGACATCTGATCCACCA 1634
Db 805 oSerAlaSerLeuThrLeuLeu----- 812
OY 1635 GGAATCACTGACTACTGCTCTCATGCTCTCTCTCTCTCA---TCCCCACCCGCCA 1691
Db 813 .....SerLeuSerSerLeuSerGlnGlnGlnSerThrArgSe 825
OY 1692 GGGTCGACGTCGCCCTCA-----AGCTCCCGGCTGCCCGGACAG 1733
Db 825 rLeuLeuArgSerAspSerAspGluLeuValValValTyrGlnMetProValSerProValG 845
OY 1734 TTCTTCAGTGTGCTTTCACACCGACGCTGAGAGCTGACAGATGCTACAT----- 1785
Db 845 uAla-----ProProLeuSerProCysThrHisAsnProLeuValAsnValArgVa 862
OY 1786 .....-ACGCCCGGTGC 1796
Db 862 GlnArgPheLysArgAspProAsnGlnSerLeuThrProThrHisValThrLeuThrTh 882
OY 1797 TGCTTCGACAGCTGTCAGATGACCTGACCCCGATGAGGGAGAGCCTTACAGCCAGCA 1856
Db 882 rProSerGlnProSerSerHisArgArgThrProSerAspGlyAlaLeuLysProGluTh 902
OY 1857 CCTGACGAGCGGTGAGATCTCGCGCTGACAGAGACCTATTTGGTCCCAAGCGCGGTGG 1916
Db 902 rLeuLeuAlaSer-ArgSerProSerSerAsnGlyLeuSerProSerProGlyAlaLeuG 922
OY 1917 AGATGTTATGCTATTTGGCTGTGAGAGAGATTTCTGAAGCCCAAGCGCGGCGAGTCATTC 1976
Db 922 LuserSerSerPheLeuPheProPhePheValProProGlnGlyMetLeuLysThrP 942
OY 1977 CGTCT 1981
Db 942 rSer 943

```

```

APPLICANT: CHIEN, David
APPLICANT: WILSON, Amy D.
APPLICANT: SMARNARAK, Anita
APPLICANT: GORVAD, Ann E.
APPLICANT: HARFALIA, April J. A.
APPLICANT: EMERLING, Brooke M.
APPLICANT: RAMKUMAR, Jayalaxmi
APPLICANT: JIN, Pei
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: MARQUIS, Joseph P.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: CHAMLA, Nandinder K.
APPLICANT: LEHR-MASON, Patricia M.
APPLICANT: KHARE, Reena
APPLICANT: LEE, Sally
APPLICANT: HARKINS, Phillip R.
APPLICANT: BECHA, Shanya D.
APPLICANT: LEE, Soo Yeun
APPLICANT: SPRAGUE, William W.
APPLICANT: ZEBARJADIAN, Yeganeh
TITLE OF INVENTION: KINASES AND PHOSPHATASES
FILE REFERENCE: PF-1315 PCT
CURRENT APPLICATION NUMBER: PCT/US02/39126
CURRENT FILING DATE: 2002-12-06
PRIOR APPLICATION NUMBER: US 60/340,235
PRIOR FILING DATE: 2001-12-07
PRIOR APPLICATION NUMBER: US 60/343,007
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/343,546
PRIOR FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US 60/354,388
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: US 60/357,675
PRIOR FILING DATE: 2002-02-15
NUMBER OF SEQ IDS: 84
SOFTWARE: PERL Program
SEQ ID NO 13
LENGTH: 1081
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 3810039CD1
PCT-US02-39126-13

Alignment Scores:
Pred. No.: 2,59e-07 Length: 1081
Score: 313.00 Matches: 202
Percent Similarity: 33.08 Conservative: 104
Best Local Similarity: 21.84 Mismatches: 318
Query Match: 7.71% Indels: 303
DB: 1 Gaps: 36

US-09-836-392-8_COPY_22_2205 (1-2184) x PCT-US02-39126-13 (1-1081)
OY 7 AGGACACCTGGCGGCGCCAGATGCGATGAGAGACTTCCGAGTCCGCGAGGAGCGCAGC 66
Db 174 ArgHisAspProAspGlnAspGlnThrIleGlnAsnValArgGlnGlnAlaLys 193
OY 67 ATGCTGACGCGCTGACAGCACCCTGATGCTGCGCTCATGCGCATGACATCCACCG 126
Db 194 LeuPheAlaMetLeuLysHisProAsnIleIleAlaLeuThrArgGlyAlaCysLeuLysGlu 213
OY 127 -----CTGCTTCGCGCGGAGCTGCGCGCGCTGACAGCGCTGACAGCAGCGCTGCTGC 180
Db 214 ProAsnLeuLysLeuValMetGlnPheAlaArgGlyGlyProLeuAsnArgValLeuSer 233
OY 181 GAGAGCGCAGAGATTCCTTATACCCCTGAGACACATGCTACACCAAAAAATA--- 237
Db 234 -----GCTTACAGATGCGCTCGCGCGCTGCGCTTACTGACGACAGAAAA----- 279
OY 238 -----GCTTACAGATGCGCTCGCGCGCTGCGCTTACTGACGACAGAAAA----- 279

```

```

Db 243 ValAsnTrpAlaValGlnIleAlaLarGlyMetAsnTyrLeuHisAspIleAlaVal 262
QY 280 AACATCATCTTGTGACCTGAAGTCGACAACTTGTGTGCTCCCTT----- 330
Db 263 ProIleIleHisArgAspLeuLysSerSerAsnIleLeuIleGlnIleLysValGluAsn 282
QY 331 ---GAGCTCAAGAGACACATCAACATCAAGCTATCTGATAGGAGATTTCGAGCAGTCA 387
Db 283 GlyAspLeuSerAsnLysIle---LeuLysIleThrAspPheGlyLeuAlaArgGlnTrp 301
QY 388 TTCATGAGAGGCGCCCTGAGCGCTGAGCGCTGCTGAGCGCCCAAGATCAG 447
Db 302 HisArgThrThrLysMetSerAlaAlaGlyThrTyrAlaTrpMetAlaProIleValIle 321
QY 448 CCGCATCTTGTATGATGAGAGAGATAGATATGCTCTGATGATGAGTGTCTAGCAG 507
Db 322 ArgAlaSerMetPheSerLysGlySerAspValTrpSerTyrGlyValLeuLeuTrpGln 341
QY 508 TTGCTGTACAGACAGCGCCCTGACTGGGCCACACAGCTCCAGATTGCC----- 558
Db 342 LeuLeuThrGlyGlnValProPheArgGlyLeuAspGlyLeuAlaValAlaTyrGlyVal 361
QY 559 -----AAGAACTGTCCAAAGGCAATCCGCGGCTTCGAGCGCGGAGAGAGTGCAG 612
Db 362 AlaMetAsnLysLeuAlaLeuProIleProSerThrCysProGlnPro----- 377
QY 613 TTCGCGGACATGACAGCGCTCATGATGAGAGTCTGGGACACTAAGCCAGAGAGCAGCC 672
Db 378 -----PheAlaLysLeuMetGlnAspCysTrpAsnProAspProHisSerArg--- 393
QY 673 CTGCGCCCTGTCCGTGGTGAAGCCAGATGAAGACCGACTTTGCCACTTCATGATGAA 732
Db 394 -----ProSerPheThrAsnIleLeuAspGln 402
QY 733 CTGTGCTGTGGAGACAGACACTTCTTCTC-ATCCCGAGGCGCAGATACACCGTGT 791
Db 403 LeuThrThrIleGlnIleSerIlePhePheGlnMetProLysAspSerPheHisCysLeu 422
QY 792 GTTTGGATGAGAAAGAGAGACTCCAGAACTACACGCT-----GGTAAACACAGA 842
Db 423 GlnAspAsnTrpLysHisGlnIleGlnIleMetPheAspGlnLeuArgAlaLysGlnLys 442
QY 843 GAA-----GGCCCTCAGTGAAGCTGCAG 866
Db 443 GlnLeuArgThrTrpGlnIleGlnIleLeuThrArgAlaAlaLeuGlnIleLysAsnGln 462
QY 867 GATGTGCTGCTGGATGAAGTGAAGTGCAGCTCCAGCTCCAGAGT----- 916
Db 463 GlnLeuLeuArgArgArgGlnIleGlnIleu-ArgLysGlnIleLysPheIleGlnAla 482
QY 917 -----CCCTGTGACAGCCAGCCAGAGCAGAA----- 946
Db 482 ArgLysLeuAsnIleIleIleHisGlnLeuCysGlnIleLysProArgValLysLysArg 502
QY 947 -----TCATCATTTACACCT 962
Db 502 YsgLysPheArgLysSerArgLeuLysLysAspLysAsnArgIleSerLeuPro 522
QY 963 CAAG-----GATGTGCGCCCTTAACACACACCCACAGAGCCT 1001
Db 522 ArgAspPheGlnHisLysPheThrValGlnAlaSerProThrMetAspLysArgLysSer 542
QY 1002 GGATATCCAGCTGTGCTGACCTGCTGTGCGCTGCTGCTTA----- 1045
Db 542 euIleAsnSerArgSerSerProAlaSerProThrIleIleLeuProArgLeuArgAla 562
QY 1045 ----- 1045
Db 562 IeGlnLeuThrProGlyLysSerLysThrTrpGlyArgSerSerValValProLysG 582
QY 1046 -----TTAAAGAATTCCT-----ACCTGTCTTACG 1073
Db 582 LysGlnGlyGlnIleGlnIleLysArgAlaProLysLysLysGlyArgThrTrpGly-Pro 601

```

```

QY 1074 GGGCT----- 1079
Db 602 GlyThrLeuGlyGlnLysGlnLeuAlaSerGlyAspGlnGlyLeuLysSerLeuValAsp 621
QY 1080 -----CGCCATGGGCTGTGGCGCTGTTCCTCCGTGGCGGGGACCCCAAGAGCAG 1133
Db 622 GlyTyrLysGlnTrpSerSerSerAlaProAsnLeuValLysGly---ArgSerSer 640
QY 1134 CTGCTCTACCTGTGCTCACAACACAGCCACAGTCCAAAGTTCAGACATCGCGGATGAGA 1193
Db 641 ProAlaLeuProGlyPheThrThrSerLeuMetLys----- 652
QY 1194 GCGACGAGACAGACCCCTACCCAGTGAAGCCATGAGAGTGTCAACAGCGCTGTGAGT 1233
Db 653 -----GlnAspGlnAspSerGlnIleProGlySerGlyLysSerArgLeu----- 667
QY 1234 CTGTACAGCAATGGCGCGGCTCTCTGATGACACTGTGCTCCCT----- 1301
Db 668 -----GlnHisSerProSerGlnSerTyrLeuCysIleProPheProArgGly 683
QY 1302 -----GAGATCTGACAGCGGCTGAGCCCTACAT-----GGCCCTCCATGCT 1346
Db 684 GlnAspGlyAspGlyProSerSerAspLysIleHisGlnIleProThrProValAsnSer 703
QY 1347 TACGTCACTGTGTGACAGCTGTGAGG-----CAGAGGAGAGAGTGTCTG 1394
Db 704 AlaThrSerThrProGlnLeuThrProThrAsnSerLeuLysArgGlyAla----- 721
QY 1395 GTGCTGATGACAAAGCCAACTCCTTGATGATGATACACTCCACCACTTACAGCTGTG 1454
Db 722 -----HisHis-----ArgArg 725
QY 1455 TGCCCGGACTCTTCGCGGT----- 1475
Db 726 CysGlnValAlaLeuLeuGlyCysGlyAlaValLeuAlaIleThrGlyLeuGlyPheAsp 745
QY 1476 -----CCCGACCGCCCTCAGAG 1493
Db 746 LeuLeuGlnAlaGlyLysCysGlnLeuLeuProLeuGlnIleProGlnProAlaArg 765
QY 1494 CATGTTCCCGTGGCGCCCTTGAGACA-----CGAAACCCCGGACAGCCAGCC 1540
Db 766 GlnGlnLysArgArgGlnIleLeuPheGlnArgSerArgProArgArgSerThr 785
QY 1541 ACAGCGCAACCAAG-----TGCTTGAAGGAGACTCA 1576
Db 786 SerProProSerArgLysLeuPheLysLysGlnIleProMetLeuLeuGlyAspPro 805
QY 1577 TCGCGAGCTGAGCATGATGATGAGAGCTGGGACAGCCAGATCCTGATCCAGCAG 1636
Db 806 Ser-----AlaSerLeuThrLeuLeuSerLeuSerSerIleSerGlnCysAsnSer 822
QY 1637 AATCACTACTG----- 1648
Db 823 ThrArgSerLeuLeuArgSerAspSerAspLysIleValTyrGlnMetProValSer 842
QY 1649 -----ACTAGCTCCAGTCTCTCTACTCT-----CATCCCGACCCC 1687
Db 843 ProValGlnAlaProProLeuSerProCysThrHisAsnProLeuValAsnValArgVal 862
QY 1688 GCGAGGCTGCGAGGCTCCCTCAAGCTCCCGAGCTCCCGAGAGTTCCTCCAGTGTGC 1747
Db 863 GlnArgPheLysArgAspProAsnGlnSerLeuThrProThrHisValThrLeuThrThr 882
QY 1748 CTTTCTCCAGGAGTGGAGAGTCAAGATGATGATACATACCCCGGCTGCTCCGACA 1807
Db 883 ProSerGlnProSerSerHisArgArgThrProSerAspGlyAlaLeuLysProGlnThr 902
QY 1808 GGTGTGACATGACCTGAGCC-----CCATGAGCGGAGAGACTTCA 1849
Db 903 LeuLeuAlaSerArgSerProSerSerAsnLysLeuSerProSerProGlyAlaGlyMet 922

```

```

OY 1850 GCCAGCAGCTGACGCGGTGAAGATCTGCGCTGACAGACCTCATTTGGGCCCCAGGC 1909
DB 923 LeuLysThrProSerProSerArgAspProGlyLupheProArgLeuProAspProAsn 942
OY 1910 GCGGTGGAGATGTTATGCTCATTTGGCCGTGAGA-----AGGATTCGTGAAG 1954
DB 943 ValValPheProProThrProArgArgTyrAsnThrGlnGlnAspSerThrLeuGlnArg 962
OY 1955 CCCAGCGGGGGCGAGATCTTCCTCTTAAAGCCCGAGACCTACCTCCGATGGGGTGC 2014
DB 963 ProLysThrLeuGlnPheLeuProArgProArgProSerAlaAsnArgGlnArg----- 980
OY 2015 TGGTGATGCTGCGGTGGGCAAGAGACACTGTTGTGCACCTTGGAATGAACA 2074
DB 981 -----LeuAspProThrPTr-----PheValSerProSerHisAlaArgSer 994
OY 2075 CAGAGTGTGCTGCGCGCTGTGAGAGGGGCGGCGGAGGAGCTTCACATTTCTACC 2134
DB 995 ThrSer-----ProAlaAsnSerSerThrGln 1004
OY 2135 AGCTCTACG 2143
DB 1005 ThrProSer 1007

RESULT 15
US-10-263-929-187
Sequence 187, Application US/10263929
GENERAL INFORMATION:
APPLICANT: Kim, Jaeseob
APPLICANT: Galant, Ron
TITLE OF INVENTION: Alzheimer's Disease Linked Genes
FILE REFERENCE: LSD-07417
CURRENT APPLICATION NUMBER: US/10/263,929
CURRENT FILING DATE: 2002-10-03
NUMBER OF SEQ ID NOS: 213
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 187
LENGTH: 954
TYPE: PRT
ORGANISM: Homo sapiens
US-10-263-929-187

Alignment Scores:
Pred. No.: 2,76e-07 Length: 954
Score: 312.50 Matches: 194
Percent Similarity: 33.54% Conservative: 77
Best Local Similarity: 24.01% Mismatches: 258
Query Match: 7,70% Indels: 281
DB: 6 Gaps: 29

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-263-929-187 (1-954)
OY 55 CAGGAGGCGGAGCATGCTGCACGCGTGCAGACACCCCTGCATGCGCGTGCATGCGCATC 114
DB 144 GlnGlnAlaArgLeuPheGlyAlaLeuGlnHisProAsnIleIleAlaLeuArgGlyAla 163
OY 115 AGCATTCACCGG-----CTGTGCTTCGCGCTGCAGACCTGCGCGCTGCAGACGCTCAAC 168
DB 164 CysLeuAsnProProHisLeuCysLeuValMetGluTyrAlaArgGlyValAlaLeuSer 183
OY 169 ACCGAGCTGCTGCAGACGCGGAGATTTCTTATACCCCTGGAGACATGCTGCACC 228
DB 184 ArgValLeuAlaGlyArgArg-----ValPro-----ProHisValLeuVal 197
OY 229 CAAAAAATGCTTACGATGCTGCGGCTGCGGCTTACCTGCACAAAGAA----- 279
DB 198 ---AsnThrAlaValGlnValAlaArgGlyMetAsnTyrLeuHisAsnAspAlaProVal 216
OY 280 AACATCATGTTTGTGACCTGAAGTGGACAAACATTTGTGTGTGCTTGCCTTGACGTCAG 339
DB 217 ProIleIleHisArgAspLeuLysSerIleAsnIleLeuIle-----LeuIleValIle 234
OY 340 GAGCAGATCAAC-----ATCAAGCTATCTGACTAGCGGATTTGAGAGCAG 384

```

```

DB 235 GlnAsnHisAsnLeuAlaAspThrValLeuLysIleThrAspPheGlyLeuAlaArgGlu 254
OY 385 TCAATTCATGAGAGCGCCCTAAGCGGTGAGGCGACCTCTGCTACAGCCCGGAGATC 444
DB 255 TrpHisLysThrThrLysMetSerAlaAlaGlyThrArgAlaMetAlaProGlnVal 274
OY 445 AGGCGCTGCATTTGATATGATGAGAGAGATGATGTTCTCTGATGGAATGCTCTAC 504
DB 275 IleArgLeuSerLeuPheSerLysSerSerAspAlaIleProSerPheGlyValLeuLeuTrp 294
OY 505 GAGTTGCTGTGCAGACAGCGGCTGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 558
DB 295 GluLeuLeuThrGlyGluValProTyrArgGluIleAspAlaLeuAlaValAlaTyrGly 314
OY 559 -----AAGAAGCTGTCCAAAGGATCCGCGGCTTCGCGGCGGCGGCGGCGGCGGAGAGT 609
DB 315 ValAlaMetAsnLysLeuThrLeuProIleProSerThrCysProGluPro----- 331
OY 610 CAGTTCGCGGAGCTGCAGAGCGCTCATGATGAGTGTGCGGACTAAGCCAGAGAACGA 669
DB 332 ---PheAlaArg-----LeuLeuGlnLysCysTrpAspProAspProHisGlyArg 347
OY 670 CCGCTGCGGCTGTGCGTGTGAGGCGGAGATGAAGGACCGACTTTGCCACCTTATAT 729
DB 348 -----ProAspPheGlySerIleLeuLys 355
OY 730 GAATGCTGTGCGGAGAGACAGACAGCTTCTTC-ATCCAGGCGGCGGAGTACACCGT 788
DB 356 ArgLeuGlnValIleGlnGlnSerAlaLeuPheGlnMetProLeuGlnSerPheHisSer 375
OY 789 GGTGTTTGGGATGAAGAAAGAGAGTCCAG----- 818
DB 376 LeuGlnGlnAspTrpLysLeuGlnIleGlnHisMetPheAspAlaLeuArgThrLysGlu 395
OY 818 ----- 818
DB 396 LysGlnLeuArgSerArgGlnGlnGlnLeuLeuArgAlaAlaGlnGlnArgPheGln 415
OY 818 ----- 818
DB 416 GlnGlnGlnLeuArgArgArgGlnGlnGlnLeuLeuAlaGlnArgGlnMetAspIleValGlu 435
OY 819 ---GAACTACAC----- 827
DB 436 ArgGlnLeuHisLeuLeuMetCysGlnLeuSerGlnGlnLysProArgValArgLysArg 455
OY 828 -----GGTGTGAACACAGAGAA 845
DB 456 LysGlnAsnPheLysArgSerArgLeuLeuLysLeuArgGlnGlySerHisIleSer 475
OY 846 GGGCTCATGAGAGTGCAGAG-----GATTCGCTG 875
DB 476 LeuProSerGlyPheGlnHisLysIleThrValGlnAlaSerProThrLeuAspLysArg 495
OY 876 CCCGTGGATGAAGTGAAGTGCAGCTGCAGGCTCCAGAGATCCCTGTGAGACAGCAGCA 935
DB 496 LysGlnSerAspGlyAlaSerProProAlaSerProSerIle----- 509
OY 936 GGACCGAAGAAATCTACATCTACACCTCAAGGCGCATGTGCC-----CTTAA 983
DB 510 -----IleProArgLeuArgAlaIleArg 517
OY 984 CACACCCCAACAGCGCTTGATCTCCAGCTGTGCTGCTACCTCTTTCGCGGCTGCTGT 1043
DB 518 LeuThrProValAspCysGlyGlySer----- 527
OY 1044 TATTAAAGAATTCCTACCTGTGTAGCGGCGGCTGCAGGAGTGTGCTGTGT 1103
DB 528 -----SerGlySerSerGlyGlySerGlyThrTrp 538
OY 1104 TCCCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1159

```

```
Db 539 SerArgly---GlyProProlLysGluGluLeuValGlyGlyLysLysGlyArg 557
QY 1160 CCACAGGTCACAGTTCAGCA-----TCGGGATGAG 1192
Db 558 ThrTrpGlyProSerSerThrLeuGlnLysGluArgValGlyGlyGluGluArgLys 577
QY 1193 ACGCAGGACACACCCCTACCGAGGAGCCATGAGAGTGTCAACAGCCGCTCTGAG 1252
Db 578 GlyLeuGlyGluGlySerLysGln-----TrpSerSerAlaProasn 592
QY 1253 TCTGGT-----ACAGCAATGGGCGGCTCCTGTGCA----- 1285
Db 593 LeuGlyLysSerProLysHisThrProLeuAlaProGlyPheAlaSerLeuasnGlnMet 612
QY 1286-----TCGACTGTGCTCCCTGGAGATCTGCAGGC 1315
Db 613 GluGluPheAlaGluAlaGluAspGlyGlySerSerValProPro----- 627
QY 1316 GGCTGAGCCCTACATGAGCCCTCCATGCTTACGTGCTGTCAGCTGAGGCA 1375
Db 628-----SerProTyrSerThrProSerTyrLeuSerValProLeuProAlaGluProSer 645
QY 1376 GAGGAGGAGGAGGCTGTGTGCTGATGACAGGCCACCTCTGTGTGATGTACACT 1435
Db 646 ProGlyAlaArgAla-----ProTyrGluProThrPro 656
QY 1436 CCACCACTACCACT----- 1451
Db 657 SerAlaPro-ProAlaArgTrpGlyHisGlyAlaArgArgCysAspLeuAlaLeu 676
QY 1452 -GTGTGCCGCTACTTCTGCGGGGTCGCCACGCCCTCAGGAGACATTTCCCGCGGC 1510
Db 676 uGlyCysAlaThrLeuLeu-GlyAlaValGlyLeuGlyAlaAspValAlaGluAlaArg 696
QY 1511 CTTTGACACGGAA----- 1524
Db 696 LaAlaAspGlyGluGluAlaArgArgTrpLeuAspGlyLeuPheProAlaGly 716
QY 1525-----CCCCCGGAGCCAGCCACAGCCGCAACCCAAAGTGCTG 1564
Db 716 rGpPheProAlaGlyLeuSerProProAlaArgProHisGlyArgArgGluAspValGly 736
QY 1565 AGGGGACTCCATCGCGAGCTGAGCATCATGATGAGAGAGCTGGGACGAGATCC 1624
Db 736 roGlyLeuGlyLeuAlaPro-----SerAlaThrL 746
QY 1625 TGATTCACACAGGAATCACTCACTACTGCTCCATGCTCTACTCTCATCCAC 1684
Db 746 euValSerLeuSerSerValSerAspCysAsnSerThrArgSerLeuLeuArgSerAsp 766
QY 1685 CCCCCAGGCTCGCCAGGTCCTCAAGCTCCCGAGCTCCCGACAGCAAGTTCTTCCAGTG 1744
Db 766 eraAspGluAlaAlaProAlaProSerProProSerProProAlaProThrProT 786
QY 1745 TGCTTCTCCACCGACTGCGAGACTCAGACATGATACATAGCCCGGCTGCTCCG 1804
Db 786 hrPro-SerProSerThrAsnProLeuValAspLeuGluLeuGluSerPheLysAsp 805
QY 1805 ACAGGCTGACATGACTGACCCCATGAGAGGAGACCTTACGACGACCTGCAGG 1864
Db 806 ProGlyLysLeuThr-----ProThrHisValThrAlaValCysAlaValSerArg 823
QY 1865 CCGTGAAGATCTCGCGCTCAGAGACCTCATTTGGGTCCCGAGCGCGGTGAGATGTTA 1924
Db 824 GlyHisArgArgThrProSerAspGlyAlaLeuGlyGlnArgGlyProProGluProAla 843
QY 1925 TCGTCATTGGCTGGA 1940
Db 844 Gly-HisGlyProGly 848
```

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: Apr11 15, 2003, 07:26:24 ; Search time 2688 Seconds
(without alignments)
6809.424 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822
Sequence: 1 MLRHRATDANKNFSEFROE.....IFYOSYEEGLRLEACTRRKR 728

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Command line parameters:
-MODEL-frame+ .p2n.model -DEV-xlh
-Q-/cgn2_1/USPTO.spool/US09836392/runat_08042003_090541_23458/app_query.fasta_1.903
-DB-Pending_Patents_NA_Main -OEMT-fastlap -SUFFIX-p2n.rnpm -MINMATCH=0.1
-LOOPT-0 -LOOPT-0 -UNITS-bits -START-1 -END-1 -MATRIX-biosum62
-TRANS-human0.cdl -LIST-45 -DOCALLIGN-200 -THR_SCORE-pct -THR_MAX-100
-THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OEMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0
-MAXLEN-200000000 -USER-US09836392_@CGN_1.1.2251_@runat_08042003_090541_23458
-NCPU-6 -ICPU-3 -NO_XLPXY -NO_MAMP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG
-DEV_TIMEOUT-120 -NARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -XGAPOP-10 -XGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : Pending_Patents_NA_Main:*

```

1: /cgn2_6/ptodata/1/pna/US06.COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06.COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07.COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US08.COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US09.COMB.seq:*

```

29: /cgn2_6/ptodata/1/pna/US097B.COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C.COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A.COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B.COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C.COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A.COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B.COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C.COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D.COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A.COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B.COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A.COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B.COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A.COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B.COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US6000.COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001.COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002.COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003.COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004.COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005.COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006.COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007.COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008.COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009.COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010.COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011.COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012.COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013.COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014.COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015.COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016.COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017.COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018.COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019.COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020.COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021.COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022.COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023.COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024.COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025.COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026.COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027.COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028.COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029.COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030.COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031.COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032.COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033.COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034.COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035.COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036.COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037.COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038.COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039.COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040.COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3822	100.0	3496	1	Sequence 8, Appl1
2	3822	100.0	3496	32	Sequence 8, Appl1
3	3822	100.0	3579	1	Sequence 37, Appl1
4	3815	99.8	3564	1	Sequence 14, Appl1
5	3728	97.5	2652	1	Sequence 8042, Ap
6	3579.5	93.7	7015	40	Sequence 5, Appl1

DB: 22 Gaps: 0
US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-833-245-1620 (1-706)
QY 67 ATGCTGACGCGCTGACACACCCCTGATCGTGGCTCATCGGATCAGATCCACCG 126
Db 1 MetLeuHISAlaLeuGlnHISProCysIleValAlaLeuIleGlyIleSerIleHISPro 20
QY 127 CTCGTGCTGCGCTGAGAGCTGCGCGCTCAGACGCTCAACACCGCTGTCGAGAAC 186
Db 21 LeucysPheAlaLeuGlnIleuAlaProLeuSerSerIleuAlaThrValLeuSerGlnAsn 40
QY 187 GCCAGAGTCTCTCTTATACCTCCGAGACACATGCTACCCCAAAATAGCCACAG 246
Db 41 AlaArgAspSerSerPheIleProLeuGlnHISMetLeuThrGlnIleAlaThrGln 60
QY 247 ATGCGCTGCGCGCTGAGCTGCTACCTGCACAAGAAAACATCATCTTCTGACCTGAGTGC 306
Db 61 IleAlaSerGlyLeuAlaIleThrLeuHISLysLysAsnIleIlePheCysAspLeuLysSer 80
QY 307 GACAACATTCGTGTGTGCTCCCTTGTAGCTCAAGAGACATCAACATCAAGCTATCTGAC 366
Db 81 AspaSnIleLeuAlaIleThrSerLeuAspValLysGlnHISLysIleAsnIleLysLeuSerAsp 100
QY 367 TACGGGATTCGAGGACATCATTCATGAGGGCGGCTAGGGGTGAGGACATCTGCGC 426
Db 101 TyrGlyIleSerArgLysPheHISLysGlnIleuAlaLeuGlyValGlnGlyThrProGly 120
QY 427 TACGAGCGCCAGAGATCAGGCGCTCGCATTTGATATGATGAGAAAGTATGATTTCTCC 486
Db 121 TyrGlnAlaIleProGlnIleArgProArgIleValIleThrAspGlnLysValAspMetPheSer 140
QY 487 TATGGAATGCGCTTACGAGTGTGCTGACAGACGCGCTGACCTGGCCACACACAG 546
Db 141 TyrGlnMetValLeuThrGlnLeuLeuSerGlyGlnArgProAlaLeuGlyHISLysGln 160
QY 547 CTCGAGATTCGCAAGAACTGCTCAAGGGCATCCGCGGTCTGGGACAGCGGAGGAA 606
Db 161 LeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGlnProGln 180
QY 607 GTGCACTTCGCGGACGCTCAGGCGCTCATGATGAGTCTGGGACATCAAGCCAGAGAG 666
Db 181 ValGlnPheArgArgLysGlnAlaLeuMetGlnCysThrPaspThrLysProGlnLys 200
QY 667 GCACCGCTGCGGCTGCTGCTGCTGAGGACGATGAGAGCCGACTTTGCCACTTCATG 726
Db 201 ArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThrPheMet 220
QY 727 TATGAATGCTGCTGAGAGACAGACGCTTCTCTCATCCAGGCGAGAGATGACAC 786
Db 221 TyrGlnLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnIleThr 240
QY 787 GTGTGTGTTGGGATGAGAAAGAGAGTCCAGAACTACACGGTGTGAACACAGAGAG 846
Db 241 ValValPheThrPaspGlyLysGlnLysSerArgAsnThrValValAsnThrGlnLys 260
QY 847 GGGCCATGAGGAGGTGAGGATGCTGCTGCGCTGGATGAGAGGAGTGGCCAGCTCAG 906
Db 261 GlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGlnLeuGln 280
QY 907 GTCCAGAGATCCCTGTGAGACAGCCAGCAGAGCAAGAAATCTACATCTACACCTCAG 966
Db 281 ValGlnArgSerLeuThrPheAlaThrGlnAspGlnLysIleThrIleThrLeuLys 300
QY 967 GGCATGTGCGCTTAAACACACCCCAACAGGCTTGATCTCCAGCTGTGCTACCTGC 1026
Db 301 GlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAlaThrCys 320
QY 1027 TTCTTGCGCGGCTGCTGATTAATAAGAAATCTACCTGCTGCTGCGGCTGCCCAT 1086
Db 321 PheLeuAlaValProValIleLysLysAsnSerThrLeuValLeuAlaGlyLeuLysAsp 340
QY 1087 GGGCTTGTGCTGTTCTCCGTGGTGGGCGGACCCCAAGAGACAGTGTCTACTG 1146

Db 341 GlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSerThrLeu 360
QY 1147 TGCTCACACACAGCCACACAGTCCAAAGTTCAGATCGCGGATGAAGACGACGACAGAC 1206
Db 361 CysSerHISThrAlaAsnArgSerLysPheSerIleAlaAspGlnAspAlaArgGlnAsn 380
QY 1207 CCTTACCCAGTGAAGGCCATGAGGTGTCACACAGCGGCTCTGAGAGTGTGTGACAGCAT 1266
Db 381 ProThrProValLysAlaMetGlnValValAsnSerGlySerGlnValThrPyrSerAsn 400
QY 1267 GGGCGGGGCTCTCTGTGATCGACGTGCTGCTCCCTGGAGATCTGCAGGCGGCTGAGACC 1326
Db 401 GlyProGlyLeuValIleAspCysAlaSerLeuGlnIleCysArgThrLeuLysPro 420
QY 1327 TACATGCGCCCTCCATGTTACGTCAGTCGTGTCACCTGAGAGGAGAGAGAG 1386
Db 421 TyrMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgGlyGln 440
QY 1387 GTCTGTGTGCTGCTGATGACAGGCGCAACTCTTGTGTGATGTACCATCCACCTAC 1446
Db 441 ValValAlaThrCysLeuAspAspLysAlaAsnSerLeuValMetThrHISerThrThr 460
QY 1447 CAGCTGTGCGCGGCTACTCTGCGCGGCTCCCGCCAGCCCGCTCAGGAGACATGTTCCCGT 1506
Db 461 GlnLeuCysAlaArgThrPheCysGlyValProSerProLeuArgAspPheProVal 480
QY 1507 GGGCGCTTGACACAGAACCCCGGACGACGACGACGACGACGACGACGACGACGACG 1566
Db 481 ArgProLeuAspThrGlnProProAlaIleSerHISThrAlaAsnProLysValProGln 500
QY 1567 GGGGACTCCATCGCGGAGCTGAGATCATATGATGAGAGAGCTGGGACGAGATCTGTG 1626
Db 501 GlyAspSerIleAlaAspValSerIleMetThrSerGlnGlnIleGlyThrGlnIleLeu 520
QY 1627 ATCCACAGAACTCACTCACTGACTACTGCTCCAGTCTCCTACTCCTCATCCACACC 1686
Db 521 IleHISGlnLysLeuThrThrAspThrPyrCysSerMetSerThrSerSerProPro 540
QY 1687 CGCCAGCTGCCAGGTCCCGCTCAAGCTCCCGACCTCCCGACGACGATTTCTCAAGTGTG 1746
Db 541 ArgGlnAlaAlaArgSerProSerSerLeuProSerProAlaSerSerSerVal 560
QY 1747 CCTTCTCCACGACTGGGAGACTCAGACATGTAACATACCGCGGTGCTCCCTCCAGC 1806
Db 561 ProPheSerThrAspCysGlnAspSerAspMetLeuHISThrProGlyAlaAlaSerAsp 580
QY 1807 AGCTGTGAGATGACCTGACCCCATGAGAGGAGGAGACTTGCAGCAGACCTGCAGGCG 1866
Db 581 ArgSerGlnHISAspLeuThrProMetAspGlyLysThrPheSerGlnHISLeuGlnAla 600
QY 1867 GTGAAGATCTCGCGCTGAGAGACTCATTTGGTCCCGAGCGCGGTGAGAGATTATC 1926
Db 601 ValLysIleLeuAlaValAlaArgAspLeuIleThrValProArgArgGlyGlyAspValIle 620
QY 1927 GTCATGTGCTGGAGAGAGATCTGAGCGCCAGCGGGCGGAGTATGCTGCTTAAAA 1986
Db 621 ValIleGlyLeuGlnLysAspSerIleAlaGlnArgGlyArgValIleAlaValLeuLys 640
QY 1987 GCCCAGAGCTGACTCCCATGAGGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046
Db 641 AlaArgGlnLeuThrProHISGlyValLeuValAspAlaAlaValAlaLysAspThr 660
QY 2047 GTTGTGTGACTTGTGAATAAGAAACACAGAGTGTGCTGCTGCTGAGGGGCTGG 2106
Db 661 ValValCysThrPheGlnAsnGlnAsnThrGlnThrCysLeuAlaValAlaThrArgGlyThr 680
QY 2107 GGGCGCAGGAGTTCGACATTTTACACAGTCTCAGAGGAGCTGGCGGCTGAGAGCT 2166
Db 681 GlyAlaArgGlnPheAspIlePheThrGlnSerThrGlnGlnLeuGlyArgLeuGlnAla 700
QY 2167 TGCACTCGCAAGAGAGG 2184

QY 1807 AGCTGTGACATACCTGACCCCGGAGGAGAGACTTACAGCCAGCACTGACAGCC 1866
 DB 581 ArgSerGlnHisAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeuGlnAla 600
 QY 1867 GTGAAGATCTCTCCCGCCGACAGACACCTCATTTGGTGTCCCGGAGGCGGCTGGAGATGTTATC 1926
 DB 601 ValLysIleLeuAlaValAspAspLeuIleThrProArgArgGlyAspValIle 620
 QY 1927 GTGATTGGCTGGAGAGATTCGTAAGCCGACGCGGCGGAGTATGCGCTTAA 1986
 DB 621 ValIleGlyLeuGlnLysAspSerGlyAlaGlnArgGlyArgValIleAlaValLeuLys 640
 QY 1987 GCCCGAGAGAGTGCATCCGATGGGCTGTGTGAGTCTCCGCTGGCAAGAGCACT 2046
 DB 641 AlaArgIleLeuThrProHisGlyValLeuValAspAlaValAlaValLysAspThr 660
 QY 2047 GTGTGTGACACTTGTGAATGAAACAGACAGAGTGTGCTGGCGGCTGGAGGCTG 2106
 DB 661 ValValLysThrPheGlnAsnGlnAsnThrGlnTrpCysLeuAlaValTrpArgGlyTrp 680
 QY 2107 GCGCGCAGGAGTTCGACATTTTCTACAGTCTTACGAGAGAGTGGCGGCTGGAGCT 2166
 DB 681 GlyAlaArgGlnPheAspIlePheThrGlnSerThrGlnLeuGlnLysArgLeuGlnAla 700
 QY 2167 TGCACTGCGAAGAGAGG 2184
 DB 701 CythrThrGlySerArgArg 706

RESULT 7

PCT-US01-11988-1621
 ; Sequence 1621, Application PC/TUS0111988
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc.
 ; TITLE OF INVENTION: Albumin Fusion Proteins
 ; FILE REFERENCE: P5546PCT
 ; CURRENT APPLICATION NUMBER: PCT/US01/11988
 ; CURRENT FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: 60/229, 358
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: 60/256, 931
 ; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/199, 384
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 267
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1621
 ; LENGTH: 706
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US01-11988-1621

Alignment Scores:

Pred. No.: 2,46-222 Length: 706
 Score: 3700.00 Matches: 705
 Percent Similarity: 99.86% Conservative: 0
 Best Local Similarity: 99.86% Mismatches: 1
 Query Match: 91.18% Indels: 0
 Gaps: 1

US-09-836-392-8_COPY_22_2205 (1-2184) x PCT-US01-11988-1621 (1-706)

QY 67 ATGCTGACAGGCTGACACACCCCTGATCGTGGCGCTCATCGGATGAGATCCACCGC 126
 DB 1 MetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIleHisPro 20
 QY 127 CTCGCTGCTGCGCTGAGACTGGCGGCTGACAGACCTCAACACCGTCTGTCCAGAAC 186
 DB 21 LeuGlySerPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSerGlnAsn 40
 QY 187 GCCAGAGATTCTTCTTATACCCCTGGAGACATGCTCACCAAAAATAGCTTACGAC 246
 DB 41 AlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAlaIleArgLys 60

QY 247 ATGCGCTGGGCGCTGGCTTACCTGACACAGAAAACATCATCTTCTGTGACCTGAAGTCG 306
 DB 61 IleAlaSerGlyLeuAlaIleLeuHisLysLysAsnIleIlePheCysAspLeuLysSer 80
 QY 307 GACACATTTGCTGGTGGTCCCTTGACGTACAGAGACATCAACATCAAGCTATCTGAC 366
 DB 81 AspAsnIleLeuValIleTrpSerLeuAspValLysGlnHisIleAsnIleLysLeuSerAsp 100
 QY 367 TACGGATTTGAGGAGAGTATCCATGAGAGGCGCCCTAGCGCTGAGGCGACTCTGCG 426
 DB 101 TyrGlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnGlyLysProGly 120
 QY 427 TACGAGGCGCCAGAGATCAGGCTCGCATTTGTAATGATGAGAAGTATGATTTTCC 486
 DB 121 TyrGlnAlaProGlnIleArgProArgIleValLysAspIleLysValAspMetPheSer 140
 QY 487 TATGGAATGGTGTCTACAGAGTGTGCTGACAGAGAGCGCCCTGACACTGGGCGACAG 546
 DB 141 TyrGlyMetValLeuThrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHisHisGln 160
 QY 547 CTCGAGATTGCCAAGAGCTGTCCAGAGGCAATCCGCGGCTTGTGGGCGACCGGAGAA 606
 DB 161 LeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnGlnProGlnGln 180
 QY 607 GTGCAGTTCGCGGACTGCGAGCGGCTCATGATGAGTGTGGGCACTAAGCCAGAGAA 666
 DB 181 ValGlnPheArgArgLeuGlnAlaLeuMetIleGlnLysTrpAspThrLysProGlnLys 200
 QY 667 CGACCGCTGGCCCTGTGGGTGGAGGAGCCAGATGAGAGCCGACTTTGTCACCTTCATG 726
 DB 201 ArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThrPheMet 220
 QY 727 TATGAAGTGTGCTGTGGAGACAGACAGCTTCTTCATCCAGGCGCGAGGATACAC 786
 DB 221 TyrGlnLeuGlyCysGlyLysGlnThrAlaPhePheSerSerGlnGlnGlnLysThr 240
 QY 787 GTGTGTTTGGGATGGAAGAAGAGAGTCCAGAACATCACCGGTGGGAACACAGAGAG 846
 DB 241 ValValPheThrPaspGlyLysGlnGlnSerArgAsnThrValValAsnThrGlnLys 260
 QY 847 GGCCTCATGAGAGTGCAGAGAGATGTGCTGCGCTGGATGGAAGTGTGACCTGCACCTCAG 906
 DB 261 GlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGlnLeuGln 280
 QY 907 GTCCAGAGATCCTGTGGAGACGCCAGGAGACCCAGAAATCTACATACCTACCTCAAG 966
 DB 281 ValGlnArgSerLeuThrPheAlaThrGlnAspGlnLysIleTyrIleTyrThrLeuLys 300
 QY 967 GGCATGTGCCCTTAACACACCCCAACAGGCTTGATCTCCAGCTGTGTCACCTGC 1026
 DB 301 GlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAlaThrCys 320
 QY 1027 TTCCTGGCGCGCTGCTTATTAATAAAGATTCCTACCTGCTTATAGCGGCGCTGCGCAT 1086
 DB 321 PheLeuAlaValAlaProValIleLysLysAsnSerThrLeuValLeuAlaGlyLeuAlaAsp 340
 QY 1087 GGGCTTGTGGCTGTGTTCCGTGGTGGCGGACCCCAAGAGACAGTCTCTTACCTG 1146
 DB 341 GlyLeuValAlaValAlaPheProValValArgGlyThrProLysAspSerCysSerThrLeu 360
 QY 1147 TGCTCAGACACAGCCACAGAGTCCAGATTCAGCATCGGGATGGAAGACGCGCAGAGAC 1206
 DB 361 CysSerHisThrIleAlaAsnArgSerLysPheSerIleAlaAspGlnAspAlaArgGlnAsn 380
 QY 1207 CCGTACCCAGTGAAGAGGCAATGAGAGTGTCAACAGCGGCTGTGAGTGTGTACAGCAAT 1266
 DB 381 ProThrProValLysAlaMetGlnValValAsnSerGlySerGlnValTrpTyrSerAsn 400
 QY 1267 GGGCGGCGCTCTTGTGATGAGTGTGCTCTGAGATCTGCAAGCGGCTGGAGCC 1326
 DB 401 GlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgArgLeuGlnPro 420


```

Db 261 G1yeuueitgluValGlnArgMetCysCysProGlyMetLysValSerCysGlnLeuGln 280
QY GTCCAGAGATCCCTGTGGACAGCCAGCAGACAGAAAATCTACATCTACACCTCAAG 966
Db 281 ValGlnArgSerLeuThrPheAlaThrGlnAspGlnLysIleTyrIleTyrThrLeuLys 300
QY 967 GGCATGTGCCCCCTTAACACACACCCCAAGGCGCTGGATCTCCAGCTGTGTCACCTGC 1026
Db 301 GlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValValThrCys 320
QY 1027 TTTCTGGCGCGCTGCTTATTAAGAAATCTACCTGGCTTACGGCGCTGCGCGAT 1086
Db 321 PheLeuAlaValProValIleLysLysAsnSerTyrLeuValIleuAlaGlyLeuAlaAsp 340
QY 1087 GGGCTGTGGCTGTGTTCCCGCTGGTGGGAGGACCCCAAGAGACAGCTGCTCACTG 1146
Db 341 GlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSerTyrLeu 360
QY 1147 TGCTCACACAGCCACACAGGTCACAGTTCAGATCGGGATGAGACGACGCGACAG 1206
Db 361 CysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArgGlnAsn 380
QY 1207 CCCATCCAGAGGAGCCATGAGGTGTGTCAACAGCGGCTGTGAGGTGTGTGACGCAAT 1266
Db 381 ProTyrProValLysAlaMetGluValAlaAsnSerGlySerGluValAlaTyrTyrSerAsn 400
QY 1267 GGGCGGGGCGCTGTGTGATGAGCTGTGCTCCCGAGATCTGAGCGGCTGGAGCGCC 1326
Db 401 GlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysAlaTyrLeuGlnIle 420
QY 1327 TACATGGCCCCCTCCATGAGTTCAGTCACTGTGACGCTGTGAGGCGAGAGGAGAG 1386
Db 421 TyrMetAlaProSerMetValThrSerValValCysSerSerGluGlyAlaGlyIleGln 440
QY 1387 GTCTGTGTGTGCTGTGATGACAAAGCCCACTCTGTGTGTGTGTGTGTGTGTGTGTGT 1446
Db 441 ValValTyrCysLeuAspLysAlaAsnSerLeuValMetTyrHisSerThrThrTyr 460
QY 1447 CAGCTGTGTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1506
Db 461 GlnLeuCysAlaAlaGlyTyrPheCysGlyValProSerProLeuAlaGlyAspMetPheProVal 480
QY 1507 CGGCCCTGTGACACGCAACCCCGGACGACGACGACGACGACGACGACGACGACGAC 1566
Db 481 ArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysValProGln 500
QY 1567 GGGGACTCCATCGCGGACGTGACGTATGTACAGTGAAGAGCTGGGACGACGACGATCTG 1626
Db 501 GlyAspSerIleAlaAspValSerIleMetTyrSerGlnGlnLeuGlyThrGlnIleLeu 520
QY 1627 ATCCACAGCAATCACTACTGACTACTGCTCATGCTCTCTACTCTCTCTCTCTCTCTCT 1686
Db 521 IleHisGlnGlnSerLeuThrAspTyrCysSerMetSerTyrSerSerSerProPro 540
QY 1687 CGCCAGGCTGCGAGGTCCTCCCAAGCTCCCAAGCTCCCAAGCTCCCAAGCTCCCAAGTGTG 1746
Db 541 ArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSerSerVal 560
QY 1747 CCTTCTCCACGAGCTGGGAGACTACACATGCTACATAGCCCGGTGCTGCTCCGAC 1806
Db 561 PropSerSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAlaSerAsp 580
QY 1807 AGGTGTGAGCATGACCTACACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1866
Db 581 ArgSerGlnHisAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeuGlnAla 600
QY 1867 GTGAGATCTCTCGCGTCAGAGACCTCATTTGGGTCCCGAGCGCGGTGTGAGATGTATC 1926
Db 601 ValLysIleLeuAlaValAlaArgAspLeuIleThrValProArgArgGlyGlyAspAlaIle 620
QY 1927 GTCATGTGCGCTGGAGAGATTTCTAAGCCAGCGGCGCGAGTCTATTCGCTTAAAA 1986

```

```

Db 621 ValIleGlyLeuGlnLysAspSerGlyAlaGlnArgGlyArgValIleAlaValLeuLys 640
QY 1987 GCCGAGAGCTGACTCCGACATGCGGCTGCTGTGATCTGCTGCTGCTGCTGCTGCTGCT 2046
Db 641 AlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLysAspThr 660
QY 2047 GTTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGCGCGCTGTGAGGCGCTG 2106
Db 661 ValValCysThrPheGlnAsnGlnAsnThrGlnTyrPheCysLeuAlaValAlaTyrArgGly 680
QY 2107 GGGCGGAGGAGTGTGACATTTCTTACAGCTACAGAGGAGGAGGAGGAGGAGGAGGAG 2166
Db 681 GlyAlaArgGlnPheAspIlePheTyrGlnSerTyrGlnGlnLeuGlyArgLeuGlnAla 700
QY 2167 TGCACTCGCAGAGAGAG 2184
Db 701 CysThrArgGlyArgArg 706

RESULT 9
US-10-132-382-6
: Sequence 6, Application US/10132382
: GENERAL INFORMATION:
: APPLICANT: WEISS, BERTRAM
: TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
: FILE REFERENCE: SCH-1811
: CURRENT APPLICATION NUMBER: US/10/132,382
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 1987
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-132-382-6

Alignment Scores:
Pred. No.: 1,966-214 Length: 1987
Score: 3574.50 Matches: 687
Percent Similarity: 94.50% Conservative: 0
Best Local Similarity: 94.50% Mismatches: 1
Query Match: 88.09% Indels: 39
Gaps: 1

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-132-382-6 (1-1987)
QY 1 ATGCTGAGGACGCTGGCGGACCGACGATGCATGAAGACTTCTCGAGTCCGCGAGAG 60
Db 1300 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 1319
QY 61 GCCAGCATGTGCACGCGCTGCAGACCCCTGCATCGTGGCGCTCATGGCATCAGCATC 120
Db 1320 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 1339
QY 121 CACCGCTCTGCTTCCGCTGAGCTGCGCGCTGCAGACGCTTCAACACGCTGTGCTC 180
Db 1340 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 1359
QY 181 GAGAACGCGACAGATTTCTCTTATACCTGGGACACATGCTCACCAAAAAATATACC 240
Db 1360 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 1379
QY 241 TACAGATGCGCTGGGCTGCGCTTACCTGCACAGAAAACATCATCTTGTGACCTG 300
Db 1380 TyrGlnIleAlaSerGlyLeuAlaTyrIleHisLysLysAsnIleIlePheCysAspLeu 1399
QY 301 AAGTGGCAACATTTGTGTGTGTGCTTCACTGCAAGAGACACATCAACATCAAGCTA 360
Db 1400 LysSerAspAsnIleLeuValThrSerLeuAspValLysGlnHisIleAsnIleLysLeu 1419
QY 361 TCTGACTACGGGATTTGAGGACGATTCATGAGGCGCGCTGAGGTGAGGAGGCACT 420
Db 1420 SerAspTyrGlyIleSerArgIleSerPheHisGlnGlyAlaLeuGlyValGlnGlyThr 1439

```


QY 1 ATGTGAGGACCTGGGGCCACCGATGCCATGAAGAATTCTCCGATTCGGGACAGAG 60
DB 1326 MetLeuArgHisLeuAlaValAlaThrHisPheAlaMetLysAsnPheSerGluPheArgGln 1345
QY 61 GCCAGCATGTGTGACGGGCTGCACACCCCTGCATCGTGGGCTCATCGGCATCAGCATC 120
DB 1346 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 1365
QY 121 CACCCGCTCTCTGCGCCCTGAGCTGCGGCGCTCAGACACCTCCACACCGCTGCTGCC 180
DB 1366 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 1385
QY 181 GAGAAAGCCAGAGATTCTCTTATACCCCTGGGACACATGCTCACCCAAAATATACC 240
DB 1386 GluAsnAlaHisAspSerSerPheIleProLeuGlnHisMetLeuThrGlnIleAla 1405
QY 241 TACCAAGATCGCCCTGGGCTGACCTGACACAGAAAACATCATCTTCTGTGACCTG 300
DB 1406 TyrGlnIleAlaSerGlyLeuAlaIleThrLeuHisLysLysAsnIleIlePheCysAspLeu 1425
QY 301 AAGTCGGACAACTTCTGTGTGTGCTTACCTGACCTCAAGGACACATCAATCAAGTA 360
DB 1426 LysSerAspAsnIleLeuValTyrSerLeuAspValLysGlnHisIleAsnIleLysLeu 1445
QY 361 TCTGACTACGGGATTTGAGGACAGTCAATTCATGAGGCGCCCTAGGCGTGAGGCGACT 420
DB 1446 SerAspTyrGlyIleSerArgIleSerPheHisGlnGlyAlaLeuGlyValGlnIleThr 1465
QY 421 CCTGGCTACGAGGCCCCAGAGATCAGGCGCTGCATGTATATGATGAGAGGTAGATAG 480
DB 1466 ProGlyTyrGlnAlaProGlnIleArgProArgIleValTyrAspGlnLysValAspMet 1485
QY 481 TTCTCTTGTGAATGTGCTCTAGAGATTGCTGTAGAGACAGCGCCCTGACACTGGGCGAC 540
DB 1486 PheSerTyrGlyMetValLeuValLeuTyrIleLeuLeuSerGlyGlnArgProAlaLeuGlnHis 1505
QY 541 CACCGAGCTCCAGATGGCCAGAGAGTGTCCAGAGGATCCGCGCGCTGTGGGAGGACCG 600
DB 1506 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnIlePro 1525
QY 601 GAGAAAGTGCAGTTCGGCGAGCTGCAGGCGCTCATGATGAGTGTGGAGCACTAAGCA 660
DB 1526 GlnGlnValGlnPheArgArgLeuGlnAlaLeuMetGlnCysTyrAspThrLysPro 1545
QY 661 GAGAAAGCCAGCTGCGCTGTGCTGTGCTGAGCCAGATGAAGAGACCCGATTTTCCAC 720
DB 1546 GlnLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 1565
QY 721 TTCATGATGAAGTGTGCTGTGGAAGAGAGAGAGTCCAGAACTTCCAGGCGGACGAG 780
DB 1566 PheMetTyrGlnLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnIleGlnGln 1585
QY 781 TACACCGTGTGTGTTTGGATGGAAGAAGAGAGTCCAGAACTTCCAGGCTGTGAACA 840
DB 1586 TyrThrValValPheThrAspGlyLysGlnGlnIleSerArgAsnTyrThrValAlaSerThr 1605
QY 841 GAGAAAGGCTTATGAGGTGCAAGAGATGTGCTGCTGGGATGAAGTGAAGTGGCCAG 900
DB 1606 GlnLysGlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGln 1625
QY 901 CTCGAGGTCACAGATCCCTGTGAGAGAGCAGCAGGAGCAGCAAAATCATCATCAACC 960
DB 1626 LeuGlnValGlnArgSerLeuThrPheAlaThrGln 1637
QY 961 CTCAAGGCGATGTCCTTAACAACACCCCAAGCGCTTGATCTCCAGCTGTGCTC 1020
DB 1637 1637
QY 1021 ACCTGCTTGTGGCGCTGCTTATTAAGAATCTCTACTGCTTCTTACGGGCTTC 1080
DB 1638 AsnSerTyrLeuValLeuAlaGlyLeu 1646

QY 1081 GCCGATGGCTGTGCTGTGTTTCCGCTGCGGGGCAACCCCAAGAGACAGCTGCTCC 1140
DB 1647 AlaAspLysLeuValAlaValPheProValValArgIleThrProLysAspSerCysSer 1666
QY 1141 TACCTGTGCTCACACACAGCCACAGAGTCCAGATTCAGCATCGGGATGAAGACGACG 1200
DB 1667 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlnAspAlaArg 1666
QY 1201 CAGAAACCCCTACCCAGAGAGGCGCATGAGGTGTCAACAGCGGCTGTGAGGTGTGATC 1260
DB 1687 GlnAsnProTyrProValLysAlaMetGlnValValAsnSerGlySerGlnValTyrPyr 1706
QY 1261 AGCAATGGGCGGCTCTCTGTGATGACTGTGCTCCCTGAGAGATCTGCAGGCGGCTG 1320
DB 1707 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgArgLeu 1726
QY 1321 GAGCCCTACATGCGCCCTCATGATGATACGTCAGTCTGTGCTGCTGAGGACGAGCG 1380
DB 1727 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgGly 1746
QY 1381 GAGAGGTCGTGTGCTGCTGATGACAGGCAACCTGCTGTGATGATACACTCCACC 1440
DB 1747 GluGlnValValTyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 1766
QY 1441 ACTTACAGCTGTGTGCGCGGTACTTGTGCGGGTCCCGACCCCTCAGGAGCATGTT 1500
DB 1767 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 1786
QY 1501 CCGGTGCGGCTTGTGAGACGGAACCCCGGACGACCCACAGGCCAACCCAAAGTG 1560
DB 1787 ProValArgProLeuAspThrGluProProAlaAsnSerHisThrAlaAsnProLysVal 1806
QY 1561 CCTGAGGGGAGCTCCATCGGCGAGGTGAGCATGATGATGAGAGAGCTGGGACGCGAG 1620
DB 1807 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGlnIleThrGln 1826
QY 1621 ATCTGATCCACAGAAATCTCATGACTACTGCTCATGCTCTCTACTCTCAATCC 1680
DB 1827 IleLeuIleHisGlnGlnIleSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 1846
QY 1681 CCACCCGCGCAGGCTGCCAGAGTCCCGCTCAAGCTCCCGACCTCCCGACGAAATTTCC 1740
DB 1847 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerProAlaSerSerSer 1866
QY 1741 AGTGTCCTTCTCCACCGAGCTGCGAGAGCTCAGACATGATCAATACGCCGCTGCC 1800
DB 1867 SerValProPheSerThrAspCysGlnAspSerAspMetLeuHisThrProGlyAlaAla 1886
QY 1801 TCCGACAGGTGTGAGCATGACCTGACCCCATGACGAGGAGACCTTCAGCCAGACCTG 1860
DB 1887 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeu 1906
QY 1861 CAGCCGCTGAAGATCTTCCGCTGAGAACTCATTTGGGTGCCAGCGCGGTGGAAT 1920
DB 1907 GlnAlaValLysIleLeuAlaValArgAspLeuIleTyrValProArgGlyValLysP 1926
QY 1921 GTTATCGTCAATGGCTGAGAGAGATCTGAAGCCAGCGGCGGCGAGTATTCCTC 1980
DB 1927 ValIleValIleGlyLeuGlnLysAspSerGlyAlaGlnArgGlyArgValIleAlaVal 1946
QY 1981 TTAAGAAGCCGAGAGCTACTCCGCGATGGGTGTGTGATGCTGCGGTGTGCAAG 2040
DB 1947 LeuLysAlaArgGlnLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 1966
QY 2041 GACACTGTGTGTGACCTTTGAATAATAAACAAGAGTGTGCTGCGCTGTGAGG 2100
DB 1967 AspThrValValCysThrPheGlnAsnGlnAsnThrGlnTyrCysLeuAlaValTyrArg 1986
QY 2101 GAGTGGGCGGACGAGGAGTGCATTTCTTCACTGCTTACAGAGAGACTGGCGGCTG 2160
DB 1987 GlyTyrGlyAlaArgGlnPheAspIlePheTyrGlnSerTyrGlnGlnLysArgLeu 2006
QY 2161 GAGGCTTGCACTCGCAAGAGA 2181

QY 1741 AGTGTGCTTTCTCCACGAGTGGAGACTCAGACATGCTACATACGCGGCTGCTGCC 1800
 |||||
 Db 1868 SerValrProheserThrapcysgluAspserThpmetLeuHnstrProglYAlaLa 1887
 |||||
 QY 1801 TCCGACAGCTGTGAGCATGACCTGACCCCATGGAGCGGGAGACCTTCAGCCACCTG 1860
 |||||
 Db 1888 SerAsprArgSerGluHnAspLeuThrProMetAspGlyGluThrPheSerGlnHnIsLeu 1907
 |||||
 QY 1861 CAGGCGCTGAAAGATCCCGCGCTGACAGACCTCATTTGGGCGCCGAGCGGGGAGAT 1920
 |||||
 Db 1908 GlnAlaValLysIleLeuAlaValArAspLeuIleThrProArGArGlyGlyAsp 1927
 |||||
 QY 1921 GTATCTCATATGCTGTGAGAAAGATTCTGAACCCAGCGGGCCGACATGATCCGCTC 1980
 |||||
 Db 1928 ValIleValIleGlyLeuGluLysAspSerGlyAlaGlnArGlyAlaValIleAlaVal 1947
 |||||
 QY 1981 TTAAGCCCGAGAGCTGACTCCGATGGGGTGTGCTGCTGCTGCCGCTGGGCAAG 2040
 |||||
 Db 1948 LeuYsAlaArGluLeuThrProHnIsGlyAlaLeuValAspAlaAlaValAlaLys 1967
 |||||
 QY 2041 GACACTGTGTGCTGACCTTGAATAAGAAACAGAGTGGCTGCGCGCTGGAGG 2100
 |||||
 Db 1968 AspThrValValCysThrPheGluAsnGluAsnThrGluTrpCysLeuAlaValTrpArg 1987
 |||||
 QY 2101 GGCTGGGGCCGAGGAGTTCGACATTTTCTACAGTCTCTACGAGAGAGTGGCGGCTG 2160
 |||||
 Db 1988 GlyTrpGlyAlaArGluPheAspIlePheThrGlnSerThrGluGluLeuGlyArgLeu 2007
 |||||
 QY 2161 GAGGCTTGACACTCCGCAAGAGA 2181
 |||||
 Db 2008 GluAlaCysThrArgLysArg 2014
 |||||

RESULT 12

US-10-132-382-4
 ; Sequence 4, Application US/10133382
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 2040
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-132-382-4

Alignment scores:

Pred. No.: 1,966-214 Length: 2040
 Score: 3574.50 Matches: 687
 Percent Similarity: 94.50% Conservatative: 0
 Best Local Similarity: 94.50% Mismatches: 1
 Query Match: 88.09% Indels: 39
 DB: 25 Gaps: 1

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-132-382-4 (1-2040)

QY 1 ATGCTGAGGACACTGCGGGCCACCGATGCCATGAAGAACTTCTCGAGTTCGCGCAGAG 60
 |||||
 Db 1353 MetLeuArGlnHnAspLeuAlaThrAspAlaMetLysAsnPheserGluPheArgGlnGlu 1372
 |||||
 QY 61 GCCAGCATGCTGCAGCGCTGACAGACCCCTGCATCTGCGGGCTCATGGCATCGCATC 120
 |||||
 Db 1373 AlaSerMetLeuHnIsAlaLeuGlnHnIsProCysIleValAlaLeuIleGlyIleSerIle 1392
 |||||
 QY 121 CACCGCTCTGCTGCGCTGAGAGCTCGCGCGCTCAGACAGCTCAACAGCGTGTCTC 180
 |||||
 Db 1393 HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 1412
 |||||
 QY 181 GAGAGCGCAGAGATTCTCTTATACCCCTGGAGACATGCTCAACCAAAAAATAGCC 240
 |||||

Db 1413 GluAsnAlaArAspSerSerPheIleProLeuGlyHnIsMetLeuThrGlnLysIleAla 1432
 |||||
 QY 241 TACAGATGCGCTCGGGCCCTGACCTGACAGAAAGAAACATCATCTCTGAGCCTG 300
 |||||
 Db 1433 TyrGlnIleAlaSerGlyLeuAlaIleTyrLeuHnIsLysLysAsnIleIlePheCysAspLeu 1452
 |||||
 QY 301 AAGTCGACCAACATTTCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 |||||
 Db 1453 LysSerAspAsnIleLeuValIleTrpSerLeuAspValLysGlnHnIsLeuAlaIleLysLeu 1472
 |||||
 QY 361 TCTGACTACGGGATTTTGAGGACACTCTTCCATGAGAGGCGCCCTGAGCGTGGAGGACT 420
 |||||
 Db 1473 SerAspIleGlyIleSerArGlnSerPheHisGluIleValAlaLeuGlyValGluGlyThr 1492
 |||||
 QY 421 CCGGCTACAGAGCGCCGAGAGATCAGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 |||||
 Db 1493 ProGlyTyrGlnAlaProGluIleArgProArgIleValIleTyrAspGluLysValAspMet 1512
 |||||
 QY 481 TTCTCTATGGAATGTGTCTCTACAGCTTGTCTGACAGACAGCGCCTGCACTGGGCGCAC 540
 |||||
 Db 1513 PheSerTyrGlyMetValLeuTyrGluLeuLeuSerGlyGlnArGProAlaLeuGlyHis 1532
 |||||
 QY 541 CACCACTCCAGATGCGCAAGAGCTCTCCAGAGGCAATCCGCGCGCTTGGGCGAGCG 600
 |||||
 Db 1533 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGlnPro 1552
 |||||
 QY 601 GAGGAGTGCAGTCCGGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 |||||
 Db 1553 GluGluValGlnPheArGArGLeuGlnAlaLeuMetGluCysTrpAspThrLysPro 1572
 |||||
 QY 661 GAGAGGACCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 |||||
 Db 1573 GluLysArGProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThr 1592
 |||||
 QY 721 TTCTATGTAACCTGTCTGTGGGAGACAGACACCTTCTCTATCCAGGCGCAGAG 780
 |||||
 Db 1593 PheMetTyrGluLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGln 1612
 |||||
 QY 781 TACACCGTGGGTGTTGGATGATGAAGAGAGTCCAGAGTCACTACAGCTGATGAACACA 840
 |||||
 Db 1613 TyrThrValValPheTrpAspGlyLysGluGlnLeuSerArgAsnTyrThrValAlaAsnThr 1632
 |||||
 QY 841 GAGAGGCGCTCATGAGAGTGCAGAGATGTGCTGCCCTGGATGAAGTACAGCTGCCAG 900
 |||||
 Db 1633 GluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 1652
 |||||
 QY 901 CTCAGGTCCAGAGATCCCTGTGACAGACCCAGAGACCAAGAAATCTACATCTACACC 960
 |||||
 Db 1653 LeuGlnValGlnArGSerLeuTrpThrAlaThrGlu 1664
 |||||
 QY 961 CTCAGGCGCATGTGCCCTTAACACACACCCAGAGGCTTGATCTACAGTGTGCTC 1020
 |||||
 Db 1664 1664
 |||||
 QY 1021 ACCTGCTTGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 |||||
 Db 1665 1665
 |||||
 QY 1081 GCCGATGGGCTTGTGCTGTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 |||||
 Db 1674 AlaAspGlyLeuValAlaValAlaPheProValAlaArgGlyThrProLysAspSerCysSer 1693
 |||||
 QY 1141 TACCTGTGCTCAGACACAGCCAGAGGCTCAAGTTCAAGTCACTGCGGAGTGAAGAGCAGCG 1200
 |||||
 Db 1694 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 1713
 |||||
 QY 1201 CAGAACCCCTACCCAGTGAAGCCATGAGGTGTGCTCAACAGCGCTGAGAGTGTGTATC 1260
 |||||
 Db 1714 GlnAsnProTyrProValIleAlaMetGluValAlaAsnSerGlySerGluValTrpPyr 1733
 |||||
 QY 1261 AGCATATGGCGCGGCTCTCTGTCTCATGACTGTGCTCTCTGAGAGATCTGAGCGGCTG 1320
 |||||
 Db 1734 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 1753
 |||||


```

Db 349 GIUGLIVAlGlnhPheArgArgLeuGlnAlaLeuMetMetGluCysTrpAspThrLysPro 368
QY 661 GAGAGCAGCAGCCGCTGCGCTGCGTGAGCCAGATGAAGA-CCGACATTGTCAC 719
Db 369 GULYArgProLeuAlaLeuSerValValSerGlnMetLysAlaProAspPheCysHis 388
QY 720 CTTCAGTATGAACTGGCTGGGAGAGAGAGCCCTTCTCTCCAGGCGCAGGA 779
Db 389 LeuHisVal**ThrValLeuTrpGlnAlaAspSerLeuLeuLeuProGlyProGly 408
QY 780 GPACACCGTGTGTTGGGATGGAAGAGAGATCCAGAGTACAGCAGTGTGAGCA 839
Db 409 ValHisArgGlyValLeuGlyTrpLysArgGlyValGlnGlnLeuHisGlyLys-AsnH 428
QY 840 AGAGAGAGGCTCAGAGAGAGAGAGATGCTGCGCTGGATGAAGTGAAGTGGCA 899
Db 428 rGluLysGlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysG 448
QY 900 GCTCCAGGTCCAGAGATCCCTGGACAGCCAGAGCCAGAAATCTACATCAAC 959
Db 448 nLeuGlnValGlnArgSerLeuTrpThrAlaTrpGlnAspGlnLysIleTrpIleTrpH 468
QY 960 CCTCAAGGCGATGCGCCCTTAACACACCCCAACAGCCCTTGAGATCCAGCTGCT 1019
Db 468 rLeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaVal 488
QY 1020 CACCTGCTTGTGGCGGCTGCTTATTAAGAAATCTCTACCTGCTTACGGGCT 1079
Db 488 lThrCysPheLeuAlaValProValIleLysLysAsnSerTrpLeuValLeuAlaGlyLe 508
QY 1080 CGCGGAGTGGCTGTGGCTGTGTTCCCGGCGGAGCGGCGCAAGAGAGCTGCTC 1139
Db 508 uAlaAspGlyLeuValAlaValAlaPheProValAlaArgGlyThrProLysAspSerCys 528
QY 1140 CTACCTGTGCTCAGACAGCAGCAGAGTCCAAATTCTACCTGCTTACGGGCT 1199
Db 528 rTyLLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspLysAlaArg 548
QY 1200 GCGAAGACCTTACCAGATGAAGGCGATGAGGTGTCACAGCGGCTCTAGGCTG 1259
Db 548 gGlnAsnProTrpProValLysAlaMetGlnValAlaAsnSerGlySerGlnValAlaTrp 568
QY 1260 CAGCAATGGCGCGGCTCTGTGATCGACTGCGCTCGAGAGTTCGACGGCGCT 1319
Db 568 rSerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnLecLysArgArgLe 588
QY 1320 GGAAGCCTTACATGGCCCTCAGTGTAGCTAGTCTGTCAGCTGAGGCGCAGAG 1379
Db 588 uGlnProTrpMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgG 608
QY 1380 GAGAGAGGTGCTGCTGCTGCTGATGACAGAGGCGCACTCTTGATGTACCACTCC 1439
Db 608 yGlnGlnValValAlaTrpCysLeuAspAspLysAlaAsnSerLeuValMetTrpHisSer 628
QY 1440 CACTTACAGCTGTGTCGCGGCTGCTGCTGCGGAGTCCAGCGCCCTCAGGAGCAT 1499
Db 628 rThrTrpGlnLeuCysAlaArgTrpPheCysGlyValProSerProLeuArgAspMetPh 648
QY 1500 TCCCGTGGCGGCTTGGACAGGAAGCCCGGAGCCAGCAGCAGCGCAACCAAGT 1559
Db 648 eProValArgProLeuAspThrGlnProProAlaAlaSerHisThrAlaAsnProLysVa 668
QY 1560 GCGTGAAGGGGATCCTCCGCGAGCAGTGAATCATATACAGTGAAGAGAGCTGGCAG 1619
Db 668 lProGlnGlyAspSerIleAlaAspValSerIleMetTrpSerGlnGlnLeuGlyThrG 688
QY 1620 GATCGATGATCCAGAGAAACACATCACTGCTACTGCTCATGCGCCCTACACTCCATC 1679
Db 688 nIleLeuIleHisGlnLeuSerLeuThrAspTrpCysSerMetSerSerTrpSerSerSe 708
QY 1680 CCGACCCCGCAGGCTGCCAGGTCCCGCTCAGCGCTCCCGAGCTCCCGCAGAGTTCTC 1739

```

```

Db 708 rProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSer 728
QY 1740 CAGTGGCTTCTCTCCAGGACTGGAGAGCTCAGACATGATACATACCCCGGCTGCTC 1799
Db 728 rSerValProPheSerThrAspCysGlnAspSerAspMetLeuHisThrProGlyAlaAl 748
QY 1800 CTCCGACAGGTCTGAGCATGACCTGACCCCAATGACGGGAGACCTTCAGCAGCAGCT 1859
Db 748 aSerAspArgSerGlnHisAspLeuThrProMetAspGlyGlnThrPheSerGlnGlnIle 768
QY 1860 GCAGGCC--GTGAAGATCCCTGCGCTCAGAGACTCATTTGGGTCCCGAGCGCGGTG 1916
Db 768 uAlaGlyArgGlnGlu-SerSerProSerGlnThrSerPheGlySerProGlyAlaValG 788
QY 1917 ACATGTTATCGTCATTTGGCTGGAGAGAGATTTCTGAAGCCAGCGGGGCGGACTATTC 1976
Db 788 lMetLeuSerSerLeuAlaTrpArgArgIleLeuAlaProSerGlyAlaGlnSerLeuP 808
QY 1977 CG 1978
Db 808 TO 808

```

RESULT 14
PCT-US02-05109-982
Sequence 982, Application PC/TUS0205109

GENERAL INFORMATION:

APPLICANT: HySeq Inc.
APPLICANT: Yang, Tom Y.
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aiqiong J.
APPLICANT: Yang, Yonghong
APPLICANT: Ma, Yungling
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Ghosh, Malabika
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrui
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 212/2-094-061
CURRENT APPLICATION NUMBER: PCT/US02/05109
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 09/810,173
PRIOR FILING DATE: 2001-03-15
NUMBER OF SEQ ID NOS: 1052
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 982
LENGTH: 320
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-05109-982

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3,21e-96	320	1679,00	319	0	0	0
Percent Similarity:		99,69%				
Best Local Similarity:		99,69%				
Query Match:		41,38%				

US-09-836-392-8_COPY_22_2205 (1-2184) x PCT-US02-05109-982 (1-320)

```

QY 1225 ATGAGAGTGTCAACAGCGGCTGTGAGGTCTGTACAGCATGGCGGCGCTTGTTC 1284
Db 1 MetGlnValValAsnSerGlySerGlnValAlaTrpTrpLysSerAsnGlyProGlyLeuVal 20

```

```

OY 1285 ATGACATGTCCTCCCTGAGATCTGCAGGCGGCTGAGAGCCCTACATGGCCCCCTCCATG 1344
DB 21 ILeaspysAlaSerLeuGluIleCysArgThrLeuGluProIlyrMetAlaProSerMet 40
OY 1345 GTTACGTACGTGTGTGACGTCTGAGGCGCAGAGGAGAGAGTCTGTGTGCTGGAT 1404
DB 41 ValThrSerValValCysSerSerGluGlyArgGlyGluGluValValTTPCysLeuAsp 60
OY 1405 GACAGGCGCAACCTCTGTGTGATGTACCTCCACCACTTACAGCTGTGTGCTGGCTAC 1464
DB 61 AspYpsAlaSerLeuValMetYrHisSerThrThrYrGluLeuCysAlaArgTyr 80
OY 1465 TTTCGCGGGTCCCGCCAGCCCTCAGGAGATGTTCCCGGGGCGCCCTGGACAGGAA 1524
DB 81 PheGlySerValProSerProLeuArgAspMetPheProValArgProLeuAspThrGlu 100
OY 1525 CCCCCGACCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 1584
DB 101 ProProAlaAlaSerHisThrAlaAsnProLysValProGluGluYAspSerIleAlaAsp 120
OY 1585 GTGAGCATCATGTACAGTGTGAGAGCTGGGACAGCAGATCTGTATCCACAGAAATCACTC 1644
DB 121 ValSerIleMetYrSerGluGluLeuGlyThrGluIleLeuIleHisGlnGluSerLeu 140
OY 1645 ACTGACTACTGCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1704
DB 141 ThrAspYrCysSerMetSerSerYrSerSerSerProAlaGlnAlaAlaArgSer 160
OY 1705 CCCTCAAGCTCCCGACAGCTCCCGACAGATGTTCTTCAGTGTGCTCTCTCTCCAGCTGC 1764
DB 161 ProSerSerLeuProSerSerProAlaSerSerSerSerValProPheSerThrAspCys 180
OY 1765 GAGGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 1824
DB 181 GluAspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgSerGluHisAspLeu 200
OY 1825 ACCCGCATGACGCGGAGACCTTCAGCCAGCAGCTGCAGGCGCGCTGAAGATCTCCGCGTC 1884
DB 201 ThrProMetAspGlyGluThrPheSerGlnHisLeuGlnAlaValLysIleLeuAlaVal 220
OY 1885 AGAGACTCTATTGGGCTCCCGACGCGCGGTGAGATGTTATGCTCATTTGGCTGGAGAG 1944
DB 221 ArgAspLeuIleThrValProArgArgGlyAspValIleValIleGlyLeuGluLys 240
OY 1945 GATTTCGAGCCAGCGGCGGCGAGTCAATGCGCTTAAAGCCCGAGAGCTGACTCCG 2004
DB 241 AspSerGlyAlaGlnArgGlyArgValIleAlaValLeuLysAlaArgGluLeuThrPro 260
OY 2005 CATGGGCTCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2064
DB 261 HisGlyValLeuValAspAlaAlaValAlaValAlaLysAspThrValValCysThrPheGlu 280
OY 2065 AATGAAACACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2124
DB 281 ArgGluAsnThrGluThrPysLeuAlaValTTPArgGlyTTPArgLysAlaArgLysPheAsp 300
OY 2125 ATTTCATCAGTCTCAGAGAGAGCTGGGCGCGCTGAGAGCTTGCACCTGCAGAGAAAG 2184
DB 301 IlePheThrGlnSerThrGluGluGluGluGluGluGluGluGluGluGluGluGluGluG 320

```

RESULT 15

US-60-155-811-774

Sequence 774, Application US/60155811

GENERAL INFORMATION:

APPLICANT: Bonazzi, Vivien

TITLE OF INVENTION: Isolated GPCR proteins, nucleic acid

FILE REFERENCE: C1000095

CURRENT APPLICATION NUMBER: US/60/155, 811

NUMBER OF SEQ ID NOS: 1060

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 774

LENGTH: 1050
TYPE: PRT
ORGANISM: Drosophila
US-60-155-811-774

Alignment Scores:
Pred. No.: 1,83e-18
Score: 435.00
Percent Similarity: 40.56
Best Local Similarity: 25.26
Query Match: 10,728
DB: 27
Gaps: 26

US-09-836-392-8_COPY_22_2205 (1-2184) x US-60-155-811-774 (1-1050)

```

OY 4 CTGAGGACCTGGGGGCGCCAGCATGAGAACTCTCCAGATTCGGCGAGAGGCC 63
DB 397 LeuGlnHis-----SerCysValAlaIleTyrCysThrAlaArgGlnGluLeu 411
OY 64 AGCATGCTGCAGCGGCTGCAGCAGCCCTGCATCGTGTGCTCATCGCATCAGCATCCAC 123
DB 412 AlaValLeuLeuThrLeuLysHisProAlaIleValProLeuValGlyIleCysIleLys 431
OY 124 CCGCTGTGCTTCCCTGAGAGCTGCGCGCGCTCAGAGCTCAACACCGTGTGCTCCGAG 183
DB 432 ProLeuAlaLeuValLeuGluLeuAlaProLeuGlyGlyLeuAspAlaLeuLeuArgHis 451
OY 184 AACCCAGAGATCTCTCTTATACCCCTGGGACACATGCTGCCCAAAATAGCTTAC 243
DB 452 TyrArgArgSerGlyAlaHis-----MetGlyProHisThrPheGlnThrLeuValLeu 469
OY 244 CAGATGCGCTGGGCGCTGCGCTACCTGCACAGAAATATCATCTTCTGTGACCTGAG 303
DB 470 GlnAlaAlaArgAlaIleGluThrLeuHisArgArgArgIleThrLysArgAspLeuLys 489
OY 304 TCGGACAATCTGTGTGTGCTCTT-----GACGTCAAGGAG 342
DB 490 SerGlnAsnValLeuValIleThrGluLeuProGlnProHisThrGluAspSerProAlaAsn 509
OY 343 CACATCAATCAAGCTATGCTACAGGATTCAGGAGATTCAGGAGATTCAGGAGGCGC 402
DB 510 LeuValHisIleLysIleLeuAspYrGlyIleSerArgGlnThrAlaProSerGlyAla 529
OY 403 CTAGCGCTGAGGAGGACCTCTGCTACAGGCCCGCAGATC-----AGCGCTCGC 453
DB 530 LysGlyPheGlyGlyThrGluGluPheMetAlaProGluIleLeuArgYrAsnGlyLys 549
OY 454 ATTGTATGATGAGAGAGTATGATGCTTCTCTAAGAAATGCTGTACAGATGCTG 513
DB 550 GluGluThrThrGluLysValAspCysPheSerPheGlyMetPheLeuYrGluAsnIle 569
OY 514 TCAGACAGCGGCTGCATGCTGGGCGCACACAGCTCCATGATTCAGCAAGAGAGTGCAG 573
DB 570 SerLeuArgGlnProPheGluGluGln-----SerIleLysGluCysIleLeuGln 587
OY 574 GGCATCCGCGCGCTGTGGGCGAGCGGAGAGTGCAGTCCGCGA---CTGCAGGCG 630
DB 588 GlySerArgProAlaLeuThrGln---ArgGluThrGlnPheProThrCysCysLeuAsp 606
OY 631 CTCATGATGAGTGTGCTGAGACACTAAGCCAGAGAGCCGCTGCGCTGTGCTGTG 690
DB 607 LeuMetValLeuCysThrHisGluGlnProArgArgArgProThrAlaSerGlnIleVal 626
OY 691 ACCCAGATGAGAGCCGCTTTCACACTTCAGCTCAG----- 726
DB 627 SerIleLeuSerAlaProGluCysIleHisLeuLeuAspValAlaAlaMetProHisSer 646
OY 727 TATGAACTGTGCTGTGGAG---CAGACAGGCTTCTCTATCCAGGCGCAGAGATGAC 783
DB 647 GluLysIleValCysGlyValPheGlnSerLeuValGlyMetGlyAspAspGluArgYs 666
OY 784 ACCGTGTGCTTTGGGATGAGAAAGAGAGATCCAGGAATCAGCGGTGTGACACAGAG 843

```

Db	667	GlyLeuGluLeuThrPheProSerPheGlySerArgIleAspIleLeuAspCysSerPro	686
Qy	844	AAGGC---CTCATGAGGCGCAGAGATGCTGC-----	876
Db	687	SerGlySerLeuLeuGlnCysAsnSerIleSerCysSerProGlnProGlnValAlaPro	706
Qy	877	-----	CGGGG 882
Db	707	ProLysThrProLysGlnGlyAlaAsnSerAlaAlaArgSerAlaGlnArgLeuProLys	726
Qy	883	ATGAAGGTGAGTCCAGCTGCAGTCCAGAGATCCCTGTGGACAGCCAGGAGACCAG	942
Db	727	MetAsnMetLeuCysCysCysLeuValGlyGlnAlaIleTrpMetGlyAspValSerGly	746
Qy	943	AAATCTACATCTACACCCCTCAAGGGCATGTGCCCCCTTAACACACCCCAACAGCCCTG	1002
Db	747	AsnLeuHisAlaLeuIleSerThrSerThrTyraHisLeuPheSer-----TyrMetLeu	764
Qy	1003	GATACCTCCAGCTGCGCAGCACCCTCTCTTGGCCGCGGCGCTGATTAAAGATATCCAC	1062
Db	765	ASP---ProAsnIleLysSerAlaValIleSerLeuValTyrMetGlnLysIleAla---	782
Qy	1063	CTGGCTTTAAGCGGCGCTCGCCGATGGCGCTTGCGCTGCTTTCCCGGTGCGGGGACAC	1122
Db	783	ArgValAlaValGlyThrHisAsnGly-----ArgValPheLeuValAspAlaThrGln	800
Qy	1123	TCGAAGGACAGCTGCTCTCTAC-----CTGTGCTCACAC	1155
Db	801	MetProSerAsnCysAlaPheAlaGlnGlySerPheValLeuThrGlnLecCysSerGly	820
Qy	1156	ACAGCCAAAGGATGTCAGATTCACCATCGCGGAGATGAAGAC-----	1194
Db	821	PheValLeuHisAlaAlaCysSerValValAspGlyIleTyrGlnLeuTrpCysGly	840
Qy	1195	-----GCAGCGGAGAACCCCTACCA-----	1215
Db	841	GlnIleAlaGlyLysIleAsnValPheProLeuAsnGlnAsnGlyValSerGlyHisGln	860
Qy	1216	-----GTAAAGCCATGAGGTGCT	1236
Db	861	AlaLeuCysHisSerGlnGlnProAsnLeuIleGlnAspValIleValAlaArgMetCys	880
Qy	1237	AACAGCGGCTGTGAGTGTGTGTACAGCAATAGGCGCGGCTCTGTGC-----	1284
Db	881	SerAsnGlnSerHisValPheSerCysLeuTyrProGlyCysMetValTyrGlnTrpAsp	900
Qy	1285	-----ATCCAGTGTGCTCCCTCGAGATGTGC---AGG	1314
Db	901	ValIleSerLysArgIleGlnAsnLysLeuAspCysSerLysLeuLeuProCysSerGln	920
Qy	1315	CGGCTGAGCCCTACATGCGCCCCCTCCATGCTTATGCTACGTGCTGTGAGAGGC	1374
Db	921	SerLeuGlnSerIleAlaIleAspGlnHisValAsnLeuIleLysCys-----	936
Qy	1375	AGAGGGAGAGAGCGTCGTGCTGCTCGATGACACAGGCCAATCCTTGTGTATGACAC	1434
Db	937	-----GlnIleSerAlaLeuAlaHisAsn	945
Qy	1435	TCCACCACTTACACAGTGTGTGCCCGGTACTTCTGGGGGTCCCCAGCCCCCTCGAGGAC	1494
Db	946	SerGlnLeuLysTyr-----IleGlyThrIleTrpGlyCysLeuIleValAlaGlnLeuHisThr	964
Qy	1495	ATGTTTCCCGTGTG-----CGGCGCTTGGACACAGGAACCCCGGACAGCCAGCCACAG	1545
Db	965	LeuArgProIleSerValPheArgProTyrGlnAsnGlnIleLysSerIleIleThrLeu	984
Qy	1546	GCCAACCCAAAGTGCTGTAGGGGAGCATCCATCGCGGACGTAGCATATGTACAGTGAAG	1605
Db	985	SerLysAspAsnValPro-----LeuIleAlaThrIleGlyArgArgTyr---Arg	1000
Qy	1606	GAGCTGGGCGACGAGATCGATGATCAACAGGAATCACTGATGATGATGCTGATGCTAC	1665
Db	1001	SerLeuIleSerArgTyrValAspSerAlaGlnSerSerThr-----LysSer	1016

QY 1666 TCTACTCTCTATCCCCACCCCGCCAGGCTGCCAGGTC 1704
 ||| |||::: ||| |||::: |||
 Db 1017 SeralValserThrProThrHisGlyAlaAlaLysSer 1029

Search completed: April 15, 2003, 07:21:42
Job time : 227 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - protein search, using frame_plus.n2p model

Run on: April 15, 2003, 07:12:52 ; Search time 44 Seconds

(without alignments) 15469.100 Million cell updates/sec

Title: US-09-836-392-8_COPY_22_2205

Perfect score: 4058

Sequence: 1 atgcctgagcaccctgcggcgc.....cttcgactcgcaagaag 2184

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 704857 seqs, 155824263 residues

Total number of hits satisfying chosen parameters: 1409714

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=tlh

-O=/cgn2.1/USOTO.spool/US09836392/runat_08042003_090330_22891/app.query.fasta_1.2375

-DB=Pending_Patents_AA_New -QFMT=fastan -SUFFIX=n2p.rapn -MIMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62

-TRANS=human0.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pcr -THR_MAX=100

-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -USER=US09836392 -RCGN 1.1 59 -runat_08042003_090330_22891

-NCPU=6 -ICPU=3 -NO_XLPRY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2.6/pctoddata/2/paa/PCT_NEW.COMB.pcp.*

2: /cgn2.6/pctoddata/2/paa/US06_NEW.COMB.pcp.*

3: /cgn2.6/pctoddata/2/paa/US07_NEW.COMB.pcp.*

4: /cgn2.6/pctoddata/2/paa/US08_NEW.COMB.pcp.*

5: /cgn2.6/pctoddata/2/paa/US09_NEW.COMB.pcp.*

6: /cgn2.6/pctoddata/2/paa/US10_NEW.COMB.pcp.*

7: /cgn2.6/pctoddata/2/paa/US60_NEW.COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3331	82.1	659	1	PCT-US03-01363-170 Sequence 170, App
2	2623	64.6	501	6	US-10-094-749-1689 Sequence 1689, App
3	341	8.4	885	5	US-09-949-016-7789 Sequence 7789, App
4	337.5	8.3	847	5	US-09-949-016-6222 Sequence 6222, App
5	337.5	8.3	847	6	US-10-263-929-188 Sequence 188, App
6	337.5	8.3	847	6	US-10-263-929-187 Sequence 327, App
7	334.5	8.2	850	6	US-10-263-929-192 Sequence 192, App
8	327.5	8.1	656	6	US-10-311-034-5 Sequence 5, App1
9	317.5	7.8	953	1	PCT-US03-04816-56 Sequence 56, App1
10	317.5	7.8	953	6	US-10-369-022-56 Sequence 56, App1
11	317.5	7.8	953	7	US-60-426-964-2 Sequence 2, App1

12	315.5	7.8	1097	6	US-10-288-798-12 Sequence 12, App1
13	315.5	7.8	1097	6	US-10-362-892-12 Sequence 12, App1
14	313	7.7	1081	1	PCT-US02-3916-13 Sequence 13, App1
15	312.5	7.7	954	6	US-10-263-929-187 Sequence 187, App1
16	308	7.6	1046	6	US-10-258-106-9 Sequence 9, App1
17	302	7.4	1096	6	US-10-210-130-130 Sequence 130, App
18	301.5	7.4	1168	6	US-10-366-683-30843 Sequence 30843, A
19	301.5	7.4	1118	6	US-10-115-678-259 Sequence 259, App
20	299	7.4	1037	5	US-09-724-676-70223 Sequence 70223, A
21	299	7.4	1037	5	US-09-724-676-70223 Sequence 70223, A
22	298.5	7.4	859	7	US-60-452-680-20336 Sequence 20336, A
23	298.5	7.4	859	7	US-60-453-135-12381 Sequence 12381, A
24	298.5	7.4	859	7	US-60-453-050-12381 Sequence 12381, A
25	298.5	7.4	891	5	US-09-949-016-7798 Sequence 7798, App
26	295.5	7.3	859	6	US-10-263-929-185 Sequence 185, App
27	291	7.2	1036	1	PCT-US02-39927-2 Sequence 2, App1
28	290.5	7.2	888	6	US-10-263-929-191 Sequence 191, App
29	287	7.1	1036	1	PCT-US02-34037-2 Sequence 2, App1
30	287	7.1	1036	1	PCT-US03-02586-24 Sequence 24, App1
31	287	7.1	1036	6	US-10-263-929-189 Sequence 189, App
32	287	7.1	1036	6	US-10-354-358-24 Sequence 24, App1
33	287	7.1	1047	5	US-09-724-676-70234 Sequence 70234, A
34	287	7.1	1047	5	US-09-724-676-70239 Sequence 70239, A
35	287	7.1	1047	5	US-09-724-676-70234 Sequence 70234, A
36	287	7.1	1047	5	US-09-724-676-70239 Sequence 70239, A
37	287	7.1	1069	5	US-09-724-676-70241 Sequence 70241, A
38	287	7.1	1069	5	US-09-724-676-70241 Sequence 70241, A
39	287	7.1	1069	5	US-09-724-676-70240 Sequence 70240, A
40	287	7.1	1069	5	US-09-724-676-70241 Sequence 70241, A
41	283.5	7.0	1002	6	US-10-263-929-193 Sequence 193, App
42	283.5	7.0	1038	7	US-60-453-135-11880 Sequence 11880, A
43	283.5	7.0	1038	7	US-60-453-050-11880 Sequence 11880, A
44	283.5	7.0	1040	1	PCT-US02-39927-5 Sequence 5, App1
45	281	6.9	1176	5	US-09-724-676-70244 Sequence 70244, A

ALIGNMENTS

RESULT 1
PCT-US03-01363-170
Sequence 170, Application PC/TUS0301363
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.; JONES, Anissa L.;
APPLICANT: DAHL, Christopher R.; GIETZEN, Darryl;
APPLICANT: CHINN, Joyce; DUFOUR, Gerard E.;
APPLICANT: JACKSON, Jennifer L.; YU, Jimmy Y.;
APPLICANT: THASON, Olivia; YAP, Pierre E.;
APPLICANT: AMSEY, Stefan R.; DAM, Tam C.;
APPLICANT: LIU, Tommy F.; GERSTIN Jr., Edward H.;
APPLICANT: PERALTA, Careyana H.; LEWIS, Samantha A.;
APPLICANT: CHEN, Alice; MARAWA, Rakesh;
APPLICANT: LAN, Ruth Y.; URASHKA, Michael;
APPLICANT: R., KRISHNAM, Sreenivasa R.; KOLLURU, Vijaykumar;
APPLICANT: PANESAR, Iqbal S.
TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
FILE REFERENCE: PT-1240 PCT
CURRENT APPLICATION NUMBER: PCT/US03/01363
CURRENT FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: US 60/349,946
PRIOR FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: US 60/349,413
PRIOR FILING DATE: 2002-01-17
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PERL Program
SEQ ID NO 170
LENGTH: 659
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: LG:337452.25; ofis:2002JAN18
PCT-US03-01363-170

Alignment Scores:

Pred. No.: 1,896-150 Length: 659
Score: 3331.00 Matches: 641
Percent Similarity: 95.71% Conservative: 6
Best Local Similarity: 94.82% Mismatches: 9
Query Match: 82.08% Indels: 20
DB: 1 Gaps: 3

US-09-836-392-8_COPY_22_2205 (1-2184) x PCT-US03-01363-170 (1-659)

112 ATGACATCCACCCGCTGCTGCTGAGAGTCCGGCCGCTGAGCAGGCTTACAGACC 171
1 IleSerIleHisProLeuCysPheAlaLeuCluAlaProLeuSerSerLeuAsnThr 20
172 GTGCTGCCGAAAGCCGAGATTTCTTTATACCCCTGGGACACATGCTCAACCA 231
21 ValLeuSerGluAsnAlaArgAspSerSerPheIleProLeuGluHisMetLeuThrGln 40
232 AAATAGCTCCAGATCCGCTGCGGCTGAGCTTACCTGCACAGAAAACATCATCTTC 291
41 LysIleAlaTyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePhe 60
292 TGTGACCTGAGTGGAGACACATTCGTGTGCTGCTGCTGACGTCAAGAGACACATCAC 351
61 CysAspLeuLysSerAspAsnIleLeuValTrpSerLeuAspValLysGluHisIleAsn 80
352 ATCAAGCTATGCTACACAGGAGATTGAGAGCATTCATTCAGAGGCGCCCTAGCGCTG 411
81 IleLysLeuSerAspTyrGlyIleSerArgIleSerPheHisGluGlyAlaLeuGlyVal 100
412 GAGGCACTCCTGCTACCGAGCCGAGATCAGGCTCCGCTCATTTGATATGAGAGAG 471
101 GluGlyThrProGlyTyrGlnAlaProGluIleArgProAlaGlyIleValTyrAspGluLys 120
472 GTAGATATGTTCTCTATGAGATGGTGGCTTACGAGATGCTGTGACGAGACGCGCTGCA 531
121 ValAspMetPheSerTyrGlyMetValLeuTyrGluLeuLeuSerGlyGlnArgProAla 140
532 CTGGGCCACACGAGCTCCAGATTCGCAAGAGCTGTCCAAAGGCGATCCGCGCTGTG 591
141 LeuGlyHisHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeu 160
592 GGGCAGCGGAGAGAGTGCATGTCGGGAGCTGCAGGCGCTCATGATGAGTGGCGGAG 651
161 GlyIleProGluGluValGlnPheArgTyrGluGlnAlaLeuMetMetGluCysTrpAsp 180
652 ACTAAGCCAGAGAGCGACCGCTGCGCTGTGCTGTGACGACAGATGAAGACCGGACT 711
181 ThrLysProGluLysArgProLeuAlaLeuSerValValSerGluMetLysAspProThr 200
712 TTGCGCACCTTCATGATGAAGTGTGCTGTGGAGAGCAGACAGCCTTCTTCATCCAG 771
201 PheIleAlaThrPheMetTyrGluLeuLysCysGlyLysGlnThrAlaPheSerSerGln 220
772 GGGCAGGAGTACCGGCTGTTTGGATGAGAAAGAGAGTCCAGAACTATACAGGTG 831
221 GlyIleGluTyrThrValValPheTrpAspGlyLysGluGluSerArgAsnTyrThrVal 240
832 GTGACACAGAGAGGGGCTCATGAGAGTGCAGAGATGCTGCTCCCGGAGATGAAGGTG 891
241 ValAsnThrGluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysVal 260
892 AGCTGCCAGCTCCAGAGTGCAGAGATCCCTGTGACAGCCAGCCAGAGACAGAAAATCTAC 951
261 SerCysGlnLeuGlnValGlnArgSerLeuTrpThrAlaThrGluAspGlnLysIleTyr 280
952 ATTCACACCTCAAGGGGATGCGCCCTTAACACACCCCAAGAGCGCTTGGATATCTCA 1011
281 IleTyrThrLeuLysGlyLysCysProLeuAsnThrProGlnGlnAlaLeuAspThrPro 300
1012 GCTGTGCTACCTCTCTTGGCCGTGCTGCTTATTAAGAAATCTTACTACTGGTCTTA 1071
301 AlaValValThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeu 320

1072 GCGGGCCCTCGCCGATGGGCTGTGTGTGCTGTTTCCCGTGTGCGGGCACCCCAAGGAC 1131
321 AlaGlyLeuAlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAsp 340
1132 AGCTGCTTCCTACCTGTGCTACACACAGCCAAAGGTCCAAAGTTCACATCGCGGATGAA 1191
341 SerCysSerTyrLeuLysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlu 360
1192 GAGCAGCGGAGAACCCCTTACCCAGTGAAGGCGCTGAGAGGGGCAACAGCGGCTGTAG 1251
361 AspAlaArgGlnAsnProTyrProValLysAlaMetGluValValAsnSerGlySerGlu 380
1252 GTCTGTACAGCAATGGAGCGGCGCTCTTGTCTACGATGCTGCTGCTGAGATCTGC 1311
381 ValTrpTyrSerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCys 400
1312 AGCGGGCTGAGCCCTCATAGTGGCCCTCCATGTTATGTCACGTCAGTCGTGCTGAG 1371
401 ArgArgLeuGluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlu 420
1372 GGCAGAGGGAGAGTCTGCTGCTGCTGCTGATGACAGGCAAGCTCTGTGATGTAC 1431
421 GlyArgGlyGluGluValValTrpCysLeuAspAspLysAlaAsnSerLeuValMetTyr 440
1432 CACTCACACACTACACAGCTGTGTGCGCGGTACTTCTGCGGGGTCCCAAGCCCTCAGG 1491
441 HisSerThrThrTyrGlnLeuLysAlaArgTyrPheCysGlyValProSerProLeuArg 460
1492 GACATGTTTCCCGTGGCGGCTTGGACAGAGAACCCCGAGCAGCAGCCAGCCGCAAC 1551
461 AspMetPheProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsn 480
1552 CCAAGAGTCCCTGAGGGGAGATCCATCGCGGAGCTGAGCATCATGATGAGTGAAGAGCTG 1611
481 ProLysValProGluGluLysAspSerIleAlaAspValSerIleMetTyrSerGluLeu 500
1612 GGCACGAGATCTGATTCACACAGAAATCACTACATGATCTGCTCATGCTCTCTAC 1671
501 GlyThrGlnIleLeuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerTyr 520
1672 TCCATATCCCGACCGCGGAGGCTGCGAGGTGCGGCGCCCTCAAGCCCTCCAGTCCCGACA 1731
521 SerSerSerProProValArgGlnAlaAlaArgSerProSerSerLeuProSerProAla 540
1732 AGTCTTCCAGTGTGCTTCTCCACCGAGTGCAGAGACTCAGACATGATCATACGCCC 1791
541 SerSerSerSerValProPheSerThrAspCysGluAspSerAspMetLeuHisMetPro 560
1792 GGTGCTGCTCCGACAGGTGTGACATGACCTGACCCCATGACGGGAGACCTTGAC 1851
561 GlyAlaAlaSerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSer 580
1852 CAGACCTGCGAGCGCGTGAAGATCCTGCGCGTCAAGACCTCATTTGGTCCCGACGCGC 1911
581 GlnHisLeuGlnAlaValLysIleLeuAlaValArgAspLeuIleTyrValProArg 600
1912 GGTGAGATGTTATGCTATGGCTGAGAGAGATTCCTGAAGCCAGCGGCGCGAGTGC 1971
601 GlyIlyAspValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgVal 620
1972 ATTGCCGCTTAAAGCCCGAGACCTGACTCCGATGGGGGTGCTGTGATGCTGCCGTG 2031
621 IleAlaValLeuLysAlaArgGluLeuThrProHisGlyIleMetProValSerSerVal 640
2032 GTGGCAAGGACACTGTTGTGTGACCTTGAATAATGAACACAGAGTGTGCTGCTGCC 2091
641 -----LysValCys----- 643
2092 GTCGTGAGGGGCTGGCGCGGAGGAGTGTGACATTTCTTACAGATCC 2139
644 ---TrpAlaGlyTrpProValArg-----AspMetValTyrMetAla 656


```

RESULT 2
US-10-094-749-1689
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RIOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: OTSUKA, MOTOYUKI
APPLICANT: YOSHITAKA, TSUTOMU
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1689
LENGTH: 501
TYPE: PRK
ORGANISM: Homo sapiens
US-10-094-749-1689

Alignment Scores:
Pred. No.: 7,37e-117 Length: 501
Score: 2623.00 Matches: 501
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 64.64% Indels: 1
Gaps: 0

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-094-749-1689 (1-501)
QY 478 ATGTTCTCTATGGAATGTCCTCTACGAGTTGCTGTCAGACAGCCGCTGCACCTGGC 537
DB 1 MetheserTYrGIyMeTValLeuTYrGIuLeuLeuSerGIyGlnArGProAlaLeuGIy 20
QY 538 CACACACAGCTCCAGATTCGCAAGAAGCTGCCAAGGGATCCCGCGTCTGGGCGAG 597
DB 21 HisHisGlnLeuGlnIleAlaLysLysLeuSerLysGIyIleArGProValLeuGIyGln 40
QY 598 CCGGAGAGAGTGCAGTTCGCGGACGACGAGCGCTCATGATGAGTGTGGAGACATAAG 657
DB 41 ProGluGluValGlnPheArGArgLysLeuGlnAlaLeuMetGluCysTrpAspThrLys 60
QY 658 CCAGAGAGAGCGCGTGGCCCTGTCGTGGTGAGCCAGATGAGAGACCGGACTTTTGGC 717
DB 61 ProGluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAla 80
QY 718 ACCCTCATGATGAAGTGTGCTGTGGAGAGACAGCGCTTCTCATCCAGGGCGAG 777
DB 81 ThrPheMetTYrGluLeuCysCysGIyLysGlnThrAlaPhePheSerSerGlnGIyGln 100
QY 778 GAGTACACCGGTGTGTTGGATGAGAAAGAGAGAGTCCAGAGACTACACGGTGTGAGAC 837
DB 101 GluTYrThrValValPheTrpAspGIyLysGluLysLeuSerArgAsnTYrThrValValAsn 120
QY 838 ACAGAGAGGGGCTCATGAGAGCTCAGAGAGATGTGCTGCCCTGGAGTGAAGTGAAGTGC 897

```

```

DB 121 ThrGluLysLysLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCys 140
QY 898 CAGCTCCAGGTCCAGAGATCCCTGTGGACAGCCACAGGAGACCAAGAAATCTACATCTAC 957
DB 141 GlnLeuGlnValGlnArgSerLeuTrpThrAlaThrGlnLysPheGlnLysIleTYrIleTYr 160
QY 958 ACCCTCAAGGAGCATGTGCCCTTAAACACACCCCAACAGAGCGCTGATCTCCAGCTGTG 1017
DB 161 ThrLeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaVal 180
QY 1018 GTCACCTGCTTCTGGCGGTGCTGTTATTAAAAAGAAATCTACCTGCTTAAAGCGGG 1077
DB 181 ValThrCysPheLeuAlaValProValIleLysLysAsnSerTYrLeuValLeuAlaGly 200
QY 1078 CTGGCGATGGGCTGTGGGTGTGTTCCTCCGTGTGGGGGACCCCAAGAGACGTGC 1137
DB 201 LeuAlaAspLysLeuValAlaValAlaPheProValValArgLysThrProLysAspSerCys 220
QY 1138 TCCCTACCTGTGCTCACACACAGCCCAACAGGATCCAGTTCAGCATCGCGGATGAAGACGA 1197
DB 221 SerTYrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAla 240
QY 1198 CGCAGAGACCCCTACCCAGTGAAGGCCATGAGGTGTGTCAACAGCGGCTGTGAGCTGC 1257
DB 241 ArgGlnAsnProTYrProValLysAlaMetGluValValAsnSerGIySerGluValTrp 260
QY 1258 TACAGCAATGGGGCGGCGCTCTGTATGATGACTGTGCTCCCTCGGAGATGTGAGCGG 1317
DB 261 TYrSerAsnGIyProGlyLeuLeuValIleAspCysLysSerLeuGluLeuCysArgArg 280
QY 1318 CTGGAGCCCTACATAGGCGCCCTCCATGTTACGTACGTGTCAGCTGTGACAGGAGAGA 1377
DB 281 LeuGluProTYrMetAlaProSerMetValThrSerValValCysSerSerGluGlyArg 300
QY 1378 GGGAGAGAGTCTGTGTGTGCTGTGATGACAGAGGCCACTCTTGTGATGTACACTCC 1437
DB 301 GlyGluGluValValTrpCysLeuAspAspLysAlaAsnSerLeuValMetTYrHisSer 320
QY 1438 ACCACCTACAGAGGTGTGGCGGCTACTTCTGGCGGGTCCCGACCGCCCAAGGAGATG 1497
DB 321 ThrThrTYrGlnLeuCysAlaArgTYrPheCysGlyValProSerProLeuAlaAspMet 340
QY 1498 TTTCCCGTGGCGGCTGTGGACAGCGAAGCCCGGACGCCACAGCGCCCAACCCAAAG 1557
DB 341 PheProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLys 360
QY 1558 GTGCTGAGGGGGACCTCATGCGGAGCTGAGCATCATGTATGACATGAGAGCTGGGACG 1617
DB 361 ValProGluLysAspSerIleAlaAspValSerIleMetTYrSerGluGluLeuGIyThr 380
QY 1618 CAGATCTGTATCCACACAGGATGATCATGATGATGATGATGATGATGATGATGATGATG 1677
DB 381 GlnIleLeuIleHisGlnGlnLysLeuThrAspTYrCysSerMetSerTYrSerSer 400
QY 1678 TCCCCACCGCCGACAGGTGCGAGGTCCCGCTCAAGCTCCACAGCTCCAGCAAGTTCT 1737
DB 401 SerProProArgGlnAlaAlaArgSerProSerSerLeuProSerProAlaSerSer 420
QY 1738 TCCAGTGTGCTTCTTCACCGAGCTGCGAGACTCAGACATGCTATCATAGCCGGTGTCT 1797
DB 421 SerSerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAla 440
QY 1798 GCGTCGAGAGGTGTGAGCATGACCTACCCCATGGAGCGGGGAGACTTCAGCCAGCACAC 1857
DB 441 AlaSerAspArgSerGluHisAspLeuThrProMetAspGlyLysThrPheSerGlnHis 460
QY 1858 CTGACAGCCGTGAAGATCTCGCGGTGAGAGACCTCATTTGGGTCCCAAGCGCGGTGA 1917
DB 461 LeuGlnAlaValLysIleLeuAlaValArgAspLeuIleTrpValProArgArg-ValGI 480
QY 1918 GATGTTATGTCATGTGGCTGTGGAGAGAGATTCTGAACCCAGCGGGCGGAGTATTGCC 1977

```

Db 480 uMeLeuSerSerLeuAlaIrrPaGArGILeLeuLysProSerClyAlaIuSerLeuPr 500

QY 1978 GTCr 1981

Db 500 oSer 501

RESULT 3

```

US-09-949-016-7789
: Sequence 7789, Application US/09949016
:
: GENERAL INFORMATION:
:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001107
:
: CURRENT APPLICATION NUMBER: US/09/949,016
:
: PRIOR APPLICATION NUMBER: 60/241,755-
:
: PRIOR FILING DATE: 2000-10-20
:
: PRIOR APPLICATION NUMBER: 60/237,768
:
: PRIOR FILING DATE: 2000-10-03
:
: PRIOR APPLICATION NUMBER: 60/231,498
:
: PRIOR FILING DATE: 2000-09-08
:
: NUMBER OF SEQ ID NOS: 207012
:
: SOFTWARE: fastseq for windows version 4.0
:
: SEQ ID NO 7789
:
: LENGTH: 885
:
: TYPE: PRT
:
: ORGANISM: Human
:
: US-09-949-016-7789

```

Alignment Scores:

Pred. No.:	1,23e+08	Length:	885
Score:	341.00	Matches:	1955
Percent Similarity:	34.09%	Conservative:	75
Best Local Similarity:	24.62%	Mismatches:	2411
Query Match:	8.40%	Indels:	283
DB:	5	Gaps:	38

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-949-016-7789 (1-885)

QY	52	CGGAGGAGGCGCAGACATCTGCACAGCCCTCAGACACCCCTGCATCTGAGGCGGC	1111
Db	200	ArgingluAlaArgLeuPheAlaMetLeuAlaHisProxsnlleIleIleAlaLeuGlyAla	219
QY	112	ATCAGCATCCACCCG-----CTCTGCTTGCCCTGAGACTCGCGCGCTCAGCAGCCTC	165
Db	220	ValGlySerLeuGluProAsnLeuGlyLeuValMetGluIrrAlaAlaGlyGlyProLeu	239
QY	166	AACACCGGCGTCCGAGAACCCAGAGATTCCTCTTTATACCCCTGGAGACATCTC	225
Db	240	SerArgAlaLeuAlaGlyAlaGlyArg-----ValPro---ProHisValLeu	255
QY	226	ACCCAAATAATGACTACAGATGCGCTCGGCGCTGCCTACCTGCACAGAAA-----	279
Db	254	Val---AsnTrpAlaValGlnIleAlaArgGlyMetHisTrpLeuHisCysGluAlaLeu	272
QY	280	---AACATCATCTCTGTGACCTGAACTGGAGCAACAATTTGTGGTGG-----TCC	322
Db	273	ValProValIleHisArgAspLeuLeuSerHisnAsnIleLeuLeuLeuGlnProIleGlu	292
QY	328	CTTGACGCTCAGAGCAGCATCAACATCAAGCTATCTACTACGAGGATTCGAGGCACTCA	387
Db	293	SerAspAspMetGluHisLeuSerLeuLeuIleHisArgPheGlyLeuAlaArgGluTrp	312
QY	388	TTCATGATGGGCGCCCTAAGCGCTGAGAGGCACCTCTCGTGCAACAGGCCCAAGATAGG	447
Db	313	HisLeuThrThrGlnMetSerAlaAlaGlyThrTrpAlaTrpMetAlaProGluValIle	332
QY	448	CCCTGCATTTGATATGATGAGAAAGTAAATATGTCCTCATGAAAGTGGCTACAGG	507
Db	333	LysAlaSerThrPheSerLeuGlySerAspValTrpPheGlyValLeuLeuTrpGlu	352
QY	508	TTCGTCTCAGAGCAGCGCCCTGCACCTGGGCCACCAACCAAGCTCCAGATGCC-----	558

D	b	353	LeuLeuThrcGluValProTyrArgGlyIleAspCysLeuAlaValAlaTyrGlyVal	372
O	y	559	-----AAGAAGCTCCAAAGGCATCCGCCGGTTCTGGGGACCGAAGAGTCAG	612
D	b	373	AlaValAsnLysLeuThrIleuProIleProSerThrCysProGlnPro-----	388
O	y	613	TTCGGCGAGTCAGAGCGCTCATGTATGGAGTCTGGGACACTAACCCAGACAGCACCG	672
D	b	389	-----PheAlaGlnIleuMetAlaAspCysTrpAlaGlnIleAspProHisAlaArgG--	404
O	y	673	CTGGCCCTGCGGTGGAGGCCAGATGAAGAACCCGACTTTTGCACTCATATGAA	732
D	b	405	-----ProAspPheAlaSerTlleuGlnGln	413
O	y	733	CTGTG-CTGTGGAGACAGACACSCCTTCTTCATCTCCAGGCCACGAGTAGTACCCTGT	791
D	b	414	LeuGlnAlaLeuGlnAlaGlnValLeuAlaArgGlnMetProGlnArgSerPheHisSerMet	433
O	y	792	GTTTTGGGATGAAAAAGAGAGCTCAG-----	818
D	b	434	GlnGlnGlyTrrPrySarArgGlnIleGlnGlyLeuPheAspGlyLeuAlaGlyGlyIlys	453
O	y	818	-----	818
D	b	454	GluLeuLeuSerArgGlnGlnGlnIleuThrArgAlaAlaArgGlnGlnArgSerGlnAla	473
O	y	819	---GAATACACCGGTGGTGAACAC-----	839
D	b	474	GlnGlnIleuArgArgArgGlnHisLeuLeuAlaGlnTrpGlnIleuGlnValaPheGlnArg	493
O	y	840	-----AGAGAAGGCTCAT-----	854
D	b	494	GluLeuThrIleuLeuGlnGlnValaAspArgGlnArgProHisValaArgArgArgArg	513
O	y	855	-----GAGAGTGCAGAGATGTCTGCC	878
D	b	514	GlyThrPheLysArgSerLysLeuArgAlaArgAspGlyGlyArgGlyIleSerMetPro	533
O	y	879	TGGGATGAAGTAGTAGTG-----	896
D	b	534	LeuAspPheLysHisArgIleThrValGlnAlaSerProGlyIleuAspArgArgArgAsn	553
O	y	897	-----CCAGCTCAGGTCAGAGATCCTGT-----	922
D	b	554	ValPheGlnValaGlyProGlnLysSerProThrPheProArgPheArgAlaIleGlnIleu	573
O	y	923	-----GGACAG-----	949
D	b	574	GluProAlaGluProGlyGlnAlaTrpGlyArgGlnSerProArgArgLeuGlnAspSer	593
O	y	950	ACATCTACACCTCAAGGGCATGTGCCCTTAACACACCCCAACAGCCTTGATTACTC	1009
D	b	594	SerAsnGlyLysArgArgAlaCys-----	1111
O	y	1010	CAGCTGTCGACCTGCTCTTGCCGCTCCTATTATAAAGA-----	1054
D	b	605	GlyProSerSerProLysProGlyGlnAlaGlnAsnGlyArgArgArgSerArgMetAsp	624
O	y	1055	ATTCTCACTCG-----TTTACGGGGCTCGCCGATGGCTTGGCTGTGTTC	1105
D	b	625	GlnAlaThrTrpTyrIleuAspSerAspAspSerSerProLeuGlySerProSerThrPro	644
O	y	1106	CCGNGGTGC---GAGGACACCCCAAAGACAGAGCTGGCTCACTGTGTCACACAGCCA	1162
D	b	645	ProAlaLeuAsnGlyAsnPropArg-----ProSerLeuGlnProGlnIleuPro	661
O	y	1163	ACAAGTCCAGTTACAGATTCGCGATGAAGACAGCACGGCAGNAACCCCTACCAAGTAGAG	1222
D	b	662	LysArgProValProAlaGlnArgGlyLysSerSerGly---ThrPro-----	676
O	y	1223	CCATGAGGTGTGCACACGCGCTCTGAGGTTGTGTACAGCANTGGCCGGGCTCTTG	1282

```

Db 677 -----LysLeuIleGlnArgAlaLeuAArgGlyThrAlaLeu-----LeuAlaSerLeu 693
Oy 1283 -----TATCG 1288
Db 694 GlyLeuGlyArgAspLeuGlnProProGlyGlyProGlyArgGlnArgGlyGlnSerPro 713
Oy 1289 ACTGGGCCCTCCGGAGATCGACGAGCGGCTGGAGCCGACACATGGCCCTCCATGTTA 1348
Db 714 ThrThrProThrProThrProAlaPro--CysProThrGlnProPro--Ser 721
Oy 1349 CGTCAGCTGTGACCTCTGAGGCGACAG----- 1378
Db 722 ProLeuIleCysPheSerLeuLysThrProAspSerProProThrProAlaProLeuLeu 751
Oy 1379 -----GGAGAGAGGTGCTGTGTGCTGCTGATGACACAGCCAACT 1417
Db 752 LeuAspLeuGlyIleProValGlyGlnArgSerAlaLysSerProArgArgGlnGln 771
Oy 1418 CCTGTGTATGATACACCTCCACCACTACCACTGTGTGTGCCGTACTCTGGGGGTCC 1477
Db 772 ProArgGlyGlyThrValSerProPro-----ProGlyThrSerArgSerAla 787
Oy 1478 CCAGCCCTCCAGGACATGTTCCCGTCGCGCCCTTGACACAGGAACCCGCGACCA 1537
Db 788 ProGlyThrProGlyThr-----ProArgSerPro 797
Oy 1538 -----GCCACACGCGCCACCCAAAGGTGCTGTAGGGGGACCTCCATCGCGAGC 1585
Db 798 ProLeuGlyLeuIleSerArgPro-----ArgProSerProLeuArgSer 812
Oy 1586 TGAGCATATGATGACAGTGGAGAGCTGGGACGACGACATCGTGAATCCACACAGATCACTCA 1645
Db 813 -----ArgLleAspProThr-----Ser 818
Oy 1646 CTGACTACTGCTCACTGCTCTCTACTCTCATCCCACTCCGACCGCAGCGCCAGGTCCC 1705
Db 819 PheValSerAla-----GlyProArgProSerProLeuProSerPro 832
Oy 1706 CCTCAAGCTCCCGACGCTCCCGACGACAGTCTTCACAGTGCTTCCTTCCTCCACGACTCG 1765
Db 833 -----GlnProAlaProArgArgAla--ProThrPheLeuPheProAspSerAsp 848
Oy 1766 AGGACTCAGACATGTCACATATAGCGCGGAGTGCTCCCGACAGAGTGTGACAGATGACTGA 1825
Db 848 ----- 848
Oy 1826 CCCCATGAGCGGGGAGACT--TCAGCGACGACCTGCAGCGCGTGAAGATCTCGCCG 1882
Db 849 ProPheThrAspSerProProAlaAsn--ProPheGlnGlyGlyProGlnAspCysArgAl 868
Oy 1883 TCAGAGACCTCATTTGGGTCCCGCAGCGCGGTGG 1916
Db 868 aglThrLysAspMetGlyAlaGlnAlaProThr 879

RESULT 4
US-09-949-016-6222
; Sequence 6222, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMERPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6222

```

```

; LENGTH: 847
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6222

Alignment Scores:
Pred. No.:      1,81e-08          length:      847
Score:         337.50           Matches:     195
Percent Similarity: 34.09%       Conservative: 75
Best Local Similarity: 24.62%    Mismatches: 241
Query Match:     8.32%           Gaps:      283
DB:              5               Indels:    38

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-949-016-6222 (1-847)

QY   52 CGCGAGGAGCCAGACCTCTGCACGCCGTGCAGACCCTTCATCSTGCSTGCATCCGGC 111
    ||||| ||||| :||| ||||| ||||| ||||| ||||| |||||
Db   162 ArggingualAargLeuPhealAmethleuAlahisProsnillelialaleuLyala 1811
QY   112 ATCAGACATCCACCCG-----CTGGTTCGCCCTGGAGCGCGCGCGCACAGAGCTC 165
    ||| :||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db   182 ValCysIleugluInProasnIeucysLeuValMetglutryrAlalaglygIproLeu 201
QY   166 AACACCGCTGCTGTCGAGAACGCCAGAGATTCTTCATTAAACCCSGGACACATGCTC 225
    ||||| :||| :||| ||||| ||||| ||||| ||||| ||||| |||||
Db   202 SeratgalaleuAlaglyArgArg-----ValPro--ProHisvalLeu 215
QY   226 ACCCAAAAATAAGCTACCAAGATGGCTCGCGGCTGGCTGCTACCTGCACAAAGAA----- 279
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   216 Val---AsntrpAlaValglnIlelalarglyImethistyleuhIsyglualaLeu 234
QY   280 ---AACATCATCTTCTGTGACTGAAGTGCGACAACATTTGCTGTGTGG-----TCC 327
    :||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   235 ValProvalIlehisArgdrleuLysSerhsnIleleuleuLeuInProIleu 254
QY   328 CTTGACGTCMAAGAGCACATACATCAAGCTATCTGACTACAGGGATTGAGAGCACTCA 387
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   255 SerasparmeGluinhIsylsThrIeuLysIlethrIspReeglyLeuAlaArgglutrp 274
QY   388 TTCATGAGGGCGCCCTGAGGGCGTGGAGGCGACCTCTGCTACAGGCCCGCAGAGATCAG 447
    ||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   275 HIsLysTrhThrlImeSerlaIalaglhyrThralatrimetalaprogIuvalIle 294
QY   448 CCTCGCATTTATATGATGAGAAGGATATATGTTCTTCATATGAAATGTCCTTACAG 507
    :||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   295 LysalaserThrPheserLyselYseraspalrIpserPheglValleuLeutrpglu 314
QY   508 TTGCTGTACGACAGACGCCCTGCTGCTAGTGGCCACACACAGCTCCAGANTGCC----- 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   315 LeuleuthrgIguIvalrProtyrArgglyIleaspyscuuIalavalAlaTygIyAl 334
QY   559 -----AAGAACCTTCACAGGGCATCCGCCGGTTCGGGGGACCGGAGAAAGTGCAG 612
    ||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   335 AlaValasnLysLeuThrIleuProIleProsefThrcysProgIupro----- 350
QY   613 TTCGCGGCACTGCACGCTCTCATGTAAGTAGTGTGGGACACTAAGCAGAAAGCGACG 672
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   351 -----PhealaglIleumetlasprsystrpalaglnarPrOHIsIargArg--- 366
QY   673 CTGGCCCTGTCGGTGTGAAGCCAGATGAAGAACCCGAGCTTTGGCACCTTCATGTAA 722
    :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   367 -----ProAspphealaserIleleuIngln 375
QY   733 CTGTG-CTGTGGAGACAGACAGCCTTCTTCATCCACGGCCAGAGAGTACACCGTGT 791
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   376 IeuglualaleuGlualaglnValleuarglImetrProalgaBserPehIsSermet 395
QY   792 GTTTGGATGAGAAAGAGAGTCCAG----- 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   396 gIngluglyTrpLysArgglunIleInglnIleuPheasergIuleuarGalalalygIuls 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY   818 ----- 818
```

```

Db 416 GluLeuLeuSerArgGluGluGluLeuThrArgAlaAlaArgGluGlnArgSerGlnAla 435
QY 819 ---GAACTACAGCGGTGGTGAACAC----- 839
Db 436 GluGlnLeuThrArgArgGlnHisLeuLeuAlaGlnTrpGluLeuGluValPheGluArg 455
QY 840 -----AGAGAGGCGCTCAT----- 854
Db 456 GluLeuThrLeuLeuGlnGlnValAlaSerArgGlnTrpProHisValAlaArgArgArg 475
QY 855 -----GGAGTGCAGAGAGTGTCTGCC 878
Db 476 GlyThrPheLeuArgSerLysLeuArgAlaArgAspGlyGluArgLysSerMetPro 495
QY 879 TGGGATGAGGTGAGCTG----- 896
Db 496 LeuAspPheLysHisArgLysIleThrValGlnAlaSerProGlyLeuAspArgArgAsn 515
QY 897 -----CCAGCTCCAGGTCCAGAGATCCCTGT----- 922
Db 516 ValPheGluValGlyProGlyAspSerProThrPheProArgPheArgAlaIleGlnLeu 535
QY 923 -----GGACAG-----CCACGAGAGACCAAGAAATCT 949
Db 536 GluProAlaGluProGlyGlnAlaTrpGlyArgGlnSerProArgArgLeuGluAspSer 555
QY 950 ACATCTACACCCCTCAGGAGCATGTGCCCTTAACACACCCCAACAGGCGCTGATATCTC 1009
Db 556 SerAsnGlyLysArgArgAlaCys-----TrpAlaTrp 566
QY 1010 CAGCTGTCGTCACCTGCTTGTGGCCCTGTTATTAAAGA----- 1054
Db 567 GlyProSerSerProLysProGlyValAlaGlnAsnGlyArgArgArgSerArgMetAsp 586
QY 1055 ATTCCTACCTCG-----TCTTAGGGGCGCTCGCGGAGAGGCTGTGGCTGTCTTC 1105
Db 587 GluAlaThrTrpLysLeuAspSerAspSerSerProLeuGlySerProSerThrPro 606
QY 1106 CCGTGTGTGC---GGGAGCACCACCAAGAGACAGCTGCTCTACCTGTGTGCACACAGCA 1162
Db 607 ProAlaLeuAsnGlyAsnProProArg-----ProSerLeuGluProGlnGluPro 623
QY 1163 ACAGCTCCAGCTTACAGATCGCGATGACAGCAGCAGCAGACCCCTACCCAGTGAAG 1222
Db 624 LysArgProValProAlaGluArgLysSerSerGly---ThrPro----- 638
QY 1223 CCATGAGGTGTCACACAGCGGCTGTGAGTCTGTGATACAGACAGGCGGCGCTCTG 1282
Db 639 -----LysLeuIleGlnArgAlaLeuLeuArgGlyThrAlaLeu---LeuAlaSerLeu 655
QY 1283 -----TCATCG 1288
Db 656 GlyLeuGlyArgAspLeuGlnProProGlyLysProGlyArgLysArgLysGluSerPro 675
QY 1289 ACTGTGCTCTCCCTGAATCTCGCAGCGGCTGAGCCCTACATGCCCCCTCCATGCTTA 1348
Db 676 ThrThrProProThrProThrProAlaPro---CysProThrGluProProPro---Ser 693
QY 1349 CGTCAGTCGTGTCAGCTCTGAGGAGAG----- 1378
Db 694 ProLeuIleCysPheSerLeuLysThrProAspSerProProThrProAlaProLeuLeu 713
QY 1379 -----GGAGAGAGTCTGTGTGCTGGATGATGACAGCAACT 1417
Db 714 LeuAspLeuGlyIleProValGlyGlnArgSerAlaLysSerProArgArgGluGlu 733
QY 1418 CCTTGTGATGTACCACTCCACCACTACAGCTGTGTGCGGCTACTTCTGCGGGGCC 1477
Db 734 ProArgGlyGlyThrValSerProPro-----ProGlyThrSerArgSerAla 749
QY 1478 CCACCCCTCAGGAGCATGTTCCCGTGGCGCTTGAGACAGGAACCCCGGCGAGCA 1537
Db 750 ProGlyThrProGlyThr-----ProArgSerPro 759

```

```

QY 1538 -----GCCACAGCGCCCAACCCAAAGGTGCTGAGGGGAGCTCCAGCGAGC 1585
Db 760 ProLeuGlyLeuIleSerArgPro-----ArgProSerProLeuArgSer 774
QY 1586 TGAGCATCATGTACATGAGAGAGCTGGGACGAGATCTGATCCACAGGAATCACTCA 1645
Db 775 -----ArgIleAspProTrp-----Ser 780
QY 1646 CTGACTACTCTCTCATCTCTCTCTACTCTCATCTCCACCCCGCCAGCTGCCAGTCCC 1705
Db 781 PheValSerAla-----GlyProArgProSerProLeuProSerPro 794
QY 1706 CCTCAAGCTCCCGAGCTCCCGACAGAGTTCTTCCAGTGTGCTTCCACGAGCTGCG 1765
Db 795 -----GlnProAlaProArgArgAla---ProTrpThrLeuPheProAspSer 810
QY 1766 AGCAGTCAGACATGCTACATACGCCGCGGTGCTCCGACAGGTGACCATGACCTGA 1825
Db 810 ----- 810
QY 1826 CCCCATGAGAGGAGAGCT---TCAGCCAGACATGACAGCGCGTGAATCTCGCCG 1882
Db 811 ProPheTrpAspSerProProAlaAsn-ProPheGlnGlyProGlnAspCysArgAl 830
QY 1883 TCAGAGACCTCATTTGGTCCCGAGCGCGGTGG 1916
Db 830 aglnThrLysAspMetCylValGlnAlaProTrp 841

RESULT: 5
US-10-263-929-188
; Sequence 188, Application US/10263929
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Alzheimer's Disease Linked Genes
; FILE REFERENCE: LSD-07417
; CURRENT APPLICATION NUMBER: US/10/263,929
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 188
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-263-929-188

Alignment Scores:
Pred. No.: 1,81e-08 Length: 847
Score: 337.50 Matches: 195
Percent Similarity: 34.09 Conservative: 75
Best Local Similarity: 24.62 Mismatches: 241
Query Match: 8.32 Indels: 283
Gaps: 38

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-263-929-188 (1-847)
QY 52 CGGAGAGGAGGAGCATGTCGACAGCGCTGACAGACCCCTGATCGGCGCTCATCGGC 111
Db 162 ArgGlnGlnAlaArgLeuPheAlaMetLeuAlaHisProAsnIleIleAlaLeuSala 181
QY 112 ATCAGCATCCACCG-----CTCTGCTTGGCCCTGAGCTCGCGCGCTCAGACGCTC 165
Db 182 ValCysLeuGluGluProAsnLeuCysLeuValMetCylThrAlaAlaGlyLysProLeu 201
QY 166 AACACCGTGTCTGTCGGAAGACGACAGATTTCTTTTACCCCTGGACACATGCTC 225
Db 202 SerArgAlaLeuAlaGlyArgArg-----ValPro---ProHisValLeu 215
QY 226 ACCCAAAATAGGCTTACAGATGCGCTGGGCGCTGCGCTACCTGACAGAGAAA----- 279
Db 216 Val---AsnTrpAlaValGlnIleAlaArgLysIleThrLeuHisCysGlnAlaLeu 234

```


DB: 6 Gaps: 13

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-311-034-5 (1-656)

QY 28 GCCATGAAGAACTTCTCCGAG-----TTCGGCAGAGAGCCAGCATG 69
 |||||
 DB 33 AAlAAlVstIlePhaAsnYshIsthrSerIeuArgLeuArgGlnGluIleuVal 52
 QY 70 CTCACGGGCTGCACACCCCTGCATCGTGGCTCATCGGCATCAGCATCCACCGGCGC 129
 |||||
 DB 53 LeuCYshIstLeuHshIstProSerIeuIleSerIeuLeuAlaIleGlyIleArgProArg 72
 QY 130 TGCCTGGCCCTGGAGCTGGCGCCCTGCACAGCCTCAACACCGCTGCTGCCAGAGCC 189
 |||||
 DB 73 MetIleuValMetGluIleuAlaSerIeuSglSerIeuAspArgLeuIleGlnIleAsp 92
 QY 190 AGAGATTCTCTTATACCTTGGACATGCTCACCCCAAAATAGCTTACACAGATC 249
 |||||
 DB 93 AlAser-----LeuThrArgThrLeuGlnHshArgIleAlaLeuHshVal 107
 QY 250 GCCTCGGGCTGCCTTACCTGCATCAAGAAAACATCATCTTCTGCATCGTGAAGCGGAC 309
 |||||
 DB 108 AlaAspGlyLeuArgThrLeuHshSerAlaMetIleIleThrArgAspLeuIleProHsh 127
 QY 310 AACATTCTGTGTGTCTCCCTGTGACGTCAAGACATCAACATCACTATCTGACTAC 369
 |||||
 DB 128 AsnValIleuLeuPheThrLeuThrProAsnAlaIleIleAlaIleValIleAlaAsp 147
 QY 370 GGGATTGTGAGCACTATTCATGAGGCGCCCTGAGCGGTGAGGACATCTCTGCTAC 429
 |||||
 DB 148 GlyIleAlaGlnThrCYsCYsArgMetIleIleThrIleSglIleThrProGlyIle 167
 QY 430 CAGGCCCCAGAGATC---AGGCTCTGATATATATATATATATATATATATATATATAT 486
 |||||
 DB 168 ArgAlaProGluValAlaArgGlyAsnValIleThrAsnGlnIleAlaAspValIleSer 187
 QY 487 TATGGAATGTGCTCTACAGTGTG---TCAGACAGCGCCCTGCACCTGGC----- 537
 |||||
 DB 188 PheGlyLeuLeuLeuThrArgIleLeuThrThrGlyArgIleValIleGluIleuIle 207
 QY 538 -----CACACACAGCTCAGATTGCCAAGAGCTGTCCAAAGGATCCGCGC 585
 |||||
 DB 208 PheProAsnGluPheAspGluIleuGlnIleGlnIleGlyIleProAspProValIleSgl 227
 QY 586 GTTCGGGCGGACCGGAGAGTGCAGTCCGCGGCTGCAGCGCTCATGATGAGGATGC 645
 |||||
 DB 228 TyrGlyCysAlaPro-----TrpProMetValGluIleuIleuIleSglIle 243
 QY 646 TGGGACACTAAGCCAGAGAGACGACCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
 |||||
 DB 244 LeuIleSglIleuAsnProGlnIleuArgProThrSerAlaGlnIleuAspIleuIleuSer 263
 QY 706 CCGACTTTTGGCACCCTTATGATATGATGATGCTGTGTGGAGCGACACA-----GCC 756
 |||||
 DB 264 AlaGluIleuValCYsLeuThrArgArgIleLeuLeuProIleSglIleValIleValIle 283
 QY 757 TTCTTCTCATCCAGGCGGACAGATACACCGTGTGTTTGGAGTGA----- 804
 |||||
 DB 284 MetValAlaIleThrHshIleAsnSerArgHshAlaSerIleIleThrPleuGlyCYsGlyHshThr 303
 QY 805 AAAGAGAGCTCCAGAGACTACACGCTGTGAACACAGAAAGGCGCTCATGAGAGTG--- 861
 |||||
 DB 304 AspArgGlyGlnLeuSerPheIleuAspIleuAsnThrGlnGlyIleThrSerGlnIleVal 323
 QY 862 -----CAGAGGATGTGCTGCCCTGGATGAAGTGAGCTGCCAGCTCCAGTCCAGAGA 915
 |||||
 DB 324 AlaAspSerArgIleLeuCYsLeuAlaLeu-----ValHshLeuProValIleIle 340
 QY 916 TCCCTGTG-----ACAGCCAGCGAGACAGAAATCTACATCTACACCTCAAGGCG 969
 |||||
 DB 341 GluSerIlePleuIleValSerGlyThrGlnSerIleThrIleuValIleAsnThrGlnAsp 360
 QY 970 ATGTGCCCCCTTAAACACACCCCAACAGCGCTTGATCTCCAGCTGTGTGCTGCTGCTTC 1029

DB 361 GlyIleSglArgHshIleThrLeuGlnIleuIleuMetHshAspSer-----ValHshIleu 377
 QY 1030 TTGGCGGCTGCCCTATTATAA-----AAGATTCTTACCTGCTGTTCAGCGGCTCGCC 1083
 DB 378 TyrCYsAsnSerPheSerIeuSglIleuIleuSglIleuIleuValIleGlyIleThrAla 397
 QY 1084 GATGGCTGTGGCTGTGTT 1104
 DB 398 AspGlyIleuAlaIlePhe 404

RESULT 9
 PCT-US03-04816-56
 Sequence 56, Application PC/TUS0304816
 GENERAL INFORMATION:
 APPLICANT: Millennium Pharmaceuticals, Inc.
 APPLICANT: Rosenfeld, Julie Beth
 TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
 PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
 27410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
 95431, 22245, 2387, 16558, 55054, 16314, 1675, 9569 OR
 TITLE OF INVENTION: 13424 MOLECULES
 FILE REFERENCE: MP102-027
 CURRENT APPLICATION NUMBER: PCT/US03/04816
 PRIOR FILING DATE: 2003-02-19
 PRIOR APPLICATION NUMBER: US 60/360,495
 PRIOR FILING DATE: 2002-02-28
 PRIOR APPLICATION NUMBER: US 60/370,121
 PRIOR FILING DATE: 2002-04-04
 PRIOR APPLICATION NUMBER: US 60/373,010
 PRIOR FILING DATE: 2002-04-16
 PRIOR APPLICATION NUMBER: US 60/373,908
 PRIOR FILING DATE: 2002-04-19
 PRIOR APPLICATION NUMBER: US 60/377,717
 PRIOR FILING DATE: 2002-05-03
 PRIOR APPLICATION NUMBER: US 60/379,949
 PRIOR FILING DATE: 2002-05-13
 PRIOR APPLICATION NUMBER: US 60/382,409
 PRIOR FILING DATE: 2002-05-21
 PRIOR APPLICATION NUMBER: US 60/385,280
 PRIOR FILING DATE: 2002-06-03
 PRIOR APPLICATION NUMBER: US 60/386,879
 PRIOR FILING DATE: 2002-06-06
 PRIOR APPLICATION NUMBER: US 60/387,536
 PRIOR FILING DATE: 2002-06-10
 REMAINING PRIOR Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 64
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 56
 LENGTH: 953
 TYPE: PRT
 ORGANISM: Homo sapiens
 PCT-US03-04816-56

Alignment Scores:
 Pred. No.: 1.6e-07 Length: 953
 Score: 317.50 Matches: 200
 Percent Similarity: 31.96% Conservative: 86
 Best Local Similarity: 22.35% Mismatches: 258
 Query Match: 7.82% Indels: 351
 DB: 1 Gaps: 38

US-09-836-392-8_COPY_22_2205 (1-2184) x PCT-US03-04816-56 (1-953)

QY 55 CAGAGGCGACAGCTGACGCGCTGCAGACCCCTGCATCGTGGCGCTCATGCGCATC 114
 |||||
 DB 144 GlnGluAlaArgLeuPheGlyAlaLeuGlnHshProAsnIleIleAlaLeuArgGlyAla 163
 QY 115 AGCATCCACCG-----CTCTGCTTCCCTGCAGAGCTGCGCGGCTGCAGAGCTCAAC 168
 |||||
 DB 164 CysLeuAsnProProHshIleuCYsLeuValMetGlnIleuIleuValMetGlnIleuValMetGlnIleu 183

```

QY 1921 GTATGTCATGCGCTGGAGAGGATTCGAAAGCCAGCGGGCCGAGTCATTGGCGTC 1980
DB 641 VAILLEIALEIIEGLYEUGLULYSASPSEKLUALAEINRGILYRGLYVAILLEIALVAL 660
QY 1981 TTAAGAGCCGAGAGAGTCATCGCATCGGATGGTCTGTGATCGCTGCTGGCAAG 2040
DB 661 LEULYALALARGIULEUTHRPROHISGLYVALLLEULVALASPHALALVALVALALALYS 680
QY 2041 GACACCTGTGTGTGACCTTGTGAATGAAAACACAGAGTGTGCTGCGCTGTGGAG 2100
DB 681 ASPHTRVALVALYSTRHPNEGILUASNGILUASNTHTGILTPCYSLLEUALVALITPARG 700
QY 2101 GGGTGGGGGGCCGAGGAGTGTGACATTTTCTACACAGTCTACAGAGAGAGTGGGGCGGCTG 2160
DB 701 GLYTPGLYALALARGIULPHASPLIEPHERYGLINSERTYRGLUGILUENGLYALGLEU 720
QY 2161 GAGGCTTGACATCGCAAGAGAGAG 2184
DB 721 GLUALCYSTRHARGLYSARGARG 728

RESULT 2
US-09-836-392-21
; Sequence 21, Application US/09836392
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides, a
; FILE REFERENCE: PTO20PI
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-392-21

Alignment Scores:
Pred. No.: 5,83e-230 Length: 728
Score: 3822.00 Matches: 728
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.18% Indels: 0
Gaps: 0

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-836-392-21 (1-728)
QY 1 ATGCTGAGGCACTGCGGCGCCAGCATGGAAGAACTTCTCGAATTCGCGAGAG 60
DB 1 Metleuarghislleuarglatrthraspalametylsasnpheaserglupharglindlu 20
QY 61 GCCAGCATGCGGCGGCGCTGAGACCCCTGCATCGTGGGGGCTCATGGGATCGATC 120
DB 21 Alasermetleuhsialaleuenglhispocysilevalalaleuillleuglyleserille 40
QY 121 CACCGCTCTGCTTGGCGCTGAGCTCGCGGCTCAGCAGCTTCAACAGCGTGTCTCC 180
DB 41 HisproleucysphalaleuengluleuhalproleuserseileuasnThrValleuser 60
QY 181 GAGAAAGCCAGAGATTTCTTATACCCCTGGAGACACATGCTACCCAAAAATACCC 240
DB 61 GlusnalaalargaspersepherleproleuenglhismetleuThrGlindylsileala 80

```

```

QY 241 TACAGATGCGCTTGGCGCTTGGCTTACCTGACACAGAAAAATCATCTTGTGACTG 300
DB 81 TyrGlindlaleaserclyleuhalatryleuHlslylsbasnlllelephesasyapleu 100
QY 301 AAGTGGACACATTTGTGTGTGTGCTTGGACGTCAAGAGACATCAATCAAGCTA 360
DB 101 LysSeraspasnilleuvaltrpserleuaspvalylspslunhlsileasnilleysleu 120
QY 361 TCGATACAGCGGATTCGAGGCGAGTCTTCCATTCAGAGGGCGCCCTAGGCGGAGGCACT 420
DB 121 Serasptryelyllesertrginserphehisgluglyalaleuolylaluglylthr 140
QY 421 CCTGCTACAGGCGCCGAGAGATCAGAGCTCGCATTTGATATGATGAGAGATAGTATG 480
DB 141 ProgllytrglindlaleprogluileargproarglilevallyrhaspdluysValaspmet 160
QY 481 TTTCTCTATGGAATGTGCTCTACAGATTCTCTCAGAGACAGGCGCTGACTGGGCGAC 540
DB 161 PheasertryelyleuvalleuylrgluleuLeuaserclylglnargproalaleuuglyhis 180
QY 541 CACAGCTCCAGATTCGCAAGAGAGCTGTCCAGAGGCAATCCGCGGTTCTGGGCGAGCG 600
DB 181 HisglileuenglillealalyslysleuserlysglyleargProvalleuenglindlu 200
QY 601 GAGAAAGTGCAGTTCGCGGCGACTGACAGCGCCCTCATGATGAGTGTGAGACTAAGCCA 660
DB 201 GlulgluvalindlnehargargleuGlindlaleuemetclucystpaspthrlysp 220
QY 661 GAGAAAGCGACCCCTGCGCTGTGCTGTGAGTACGACAGATGAAGACCGGACTTTTGCACC 720
DB 221 Glulysargproleualalaleuservalalaserclimetylsaspptolrhphealatr 240
QY 721 TTGATATGATACGTGTGCTGTGGGAGACAGACGCTTCTTCATCCAGGCGGCGAGAG 780
DB 241 PheMETYRGLULEUCYSGLYLSGLINTHRAlalphehesererglindylglindlu 260
QY 781 TACACCGTGTGTTTGGATGAGAAAGAGAGTCCAGAACTACAGCGTGTGAACACA 840
DB 261 Tyrthrvalvalphetrpaspaglylglulgluserargasnryrthrvalalasnthr 280
QY 841 GAGAAAGGCGCTCATGAGAGTGCAGAGATGTGCTGCGCTGGAGTGAAGTGAAGTGCAG 900
DB 281 Glulysglileuemetgluvalglinargmetcyscysproglumetylsvalsercysgln 300
QY 901 CTCAGGTCCAGAGATCCCTGTGAGACAGCAGCAGAGACAGAAATCTACATTAACC 960
DB 301 LeuGlindlalnargserleuThrAlatrthrculaspnlulyslletrylletythr 320
QY 961 CTCAGGCGATGTGCCCTTAACACACACCCACAGGCGCTTGATCTCAGCTGTGCTC 1020
DB 321 LeuylsglymetcysproleuasnThrProglindlaleuasnThrProalalval 340
QY 1021 ACCTGCTTCTTGGCGGCGCTGCTTATTAAGAAATCTACCTGTGCTTACGGCGGCTC 1080
DB 341 Thrcysphelenualalvalprovalillelyslysasnsertryleuvalleuallaglyleu 360
QY 1081 GCGATGGGCGTGTGGGCTGTTCCTGCTGCGGAGGCGGCGGCGGCGGCGGCGGCTGCTC 1140
DB 361 AlaspglileuvalalalaleuvalpheprovalalargllythrProlysaspsercysSer 380
QY 1141 TACCTGTCTCACACACAGCCACAGAGTCAAGTTACAGATCGCGGATGAGAGCGACG 1200
DB 381 TyrlleucysserhsthrAlalasnargserlyrsheserillealaspdluaspalarg 400
QY 1201 CAGAAACCCCTACCCAGTGAAGGCCATGAGGTGTCAACAGCGGCTGAGAGTGTGTAC 1260
DB 401 GlinsnProtyrProvalylsalametcgluvalalasnsergllysergluvalaltrpyr 420
QY 1261 AGCAATGGCGGCGGCGCTGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 421 SerasnnglyProgllyleuvalleuallleaspysalaserleuGlulilecysargargleu 440
QY 1321 GAGCCCTACATGCGCCCTCCATGTTACGTACGTACGTGCTGAGGCGGAGAGG 1380

```

```

|||||
Db 441 GUPROTYRMetAlaIProSerMetValThrSerValValCysSerSerGluIleArgGly 460
OY 1361 GAGGAGGTCGTGTCGCTGGATGACAAAGCCAACTCCTTGATGATGACCACTCCACC 1440
Db 461 GUGUValValITrPCysLeuAspAspLysAlaAsnSerLeuValMetLysHisSerThr 480
OY 1441 ACCTACCAAGCTGTGTCGCGGATCTTTCGCGGGTCCCAAGCCCTCAGGACATGTTT 1500
Db 481 ThrTyrGlnLeuCysAlaIleArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
OY 1501 CCCGTCGCGCCCTTGACACAGCAACCCCGGACGACGACGACGACGACGACGACGACG 1560
Db 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
OY 1561 CCTGAGGGGGAGCTCCATGCGGAGCTGAGCATATGATACAGTAGAGAGCTGGGACGCG 1620
Db 521 ProGluGlyAspSerIleAlaAspValSerIleMetLysSerGluGluLeuGln 540
OY 1621 ATCTGATCCACCAAGGATACATCACTAGTACTGTCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 541 IleuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerLysSerSerSer 560
OY 1681 CCACCCCGCCAGGTCGACAGTCCCTCAAGCTCCCAAGCTCCCAAGCTCCCAAGTCTTCC 1740
Db 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
OY 1741 AGTGTGCTTTCTCCAGGACTGCGAGAGACTAGACATGCTACATACGCGCGCTGCTGCC 1800
Db 581 SerValProPheSerThrAspCysGlnAspSerAspMetLeuHisThrProGluAlaAla 600
OY 1801 TCCGACAGGTCGACATGACCTGACCCCAATGAGCGGGGAGAGACCTTTCAGCGACGCTG 1860
Db 601 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeu 620
OY 1861 CAGGCCGTGAAGATCCTGCGCGCTCAGAGACCTCATTTGGGCTCCCAAGCGCGGTGAGAT 1920
Db 621 GlnAlaValLysIleLeuAlaValAlaArgAspLeuIleTrpValProArgGlyGlyAsp 640
OY 1921 GTTATCGTCATGCGCTGGAGAGAAGATTTCTGAAGCCGAGCGGGCGAGTCTTGGCCGC 1980
Db 641 ValIleValIleGlyLeuGlnLysAspSerGlnAlaGlnArgGlyArgValIleAlaVal 660
OY 1981 TTAAGGCCGAGAGACTGACTCCGCATGGGGGTGTGTGATGATGCGCCCTGTGGCAAG 2040
Db 661 LeuLysAlaArgGluLeuThrProHisGlyValIleuValAspAlaAlaValAlaLys 680
OY 2041 GACACTGTGTGTGACCTTTGAAAAATGAAACACAGAGTGTGCTGCTGCTGAGAG 2100
Db 681 AspThrValIaIcysThrPheGlnAsnGlnAsnThrGlnTrpCysLeuAlaValIleTrpArg 700
OY 2101 GGCTGGGGCGCCAGGAGATTCGACATTTTTCACAGTCTTACGAGAGAGCTGGCGGCTG 2160
Db 701 GlyTrpGlyAlaArgGluPheAspIlePheTyrGlnSerTyrGlnGluLeuGlnValArgLeu 720
OY 2161 GAGCTTCACATCGCAGAAGAGAG 2184
Db 721 GluAlaCysThrArgLysArgArg 728

```

```

; LENGTH: 706
; TYPE: FRT
; ORGANISM: Homo sapiens
PCT-US01-01435-82

Alignment Scores:
Pred. No.: 8.79e-223 Length: 706
Score: 3707.00 Matches: 706
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 91.35% Indels: 0
DB: 1 Gaps: 0

US-09-836-392-8_COPY_22_2205 (1-2184) x PCT-US01-01435-82 (1-706)

OY 67 ATGTGCAGCGGCTGACACACCCCTGCATCTGCGGCTGCATCGGCATCAGCATCACC 126
Db 1 MetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIleHisPro 20
OY 127 CTCTGCTTCCGCTGGAGACTGGCGCGCTCAGACGCTCAACACCCGTGCTCGAGAAC 186
Db 21 LeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSerGluAsn 40
OY 187 GCGACAGATTCCTCTTATACCCCTGGGACACATGCTCACCAGAAATAGCCATACCAG 246
Db 41 AlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAlaTyrGln 60
OY 247 ATCGCTCGGGCTGGCTGCTTACCTGCACAGAAAACATCATCTTCTGTGACCTGAAGTCG 306
Db 61 IleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeuLysSer 80
OY 307 GACACATTCCTGCTGTGCTGCTCCCTTGCATCAAGAGACATCAACATCAACTATCTGAC 366
Db 81 AspaHisIleLeuValITrSerLeuAspValLysGlnHisIleAsnIleLysLeuSerAsp 100
OY 367 TACGCGATTCGAGCAGCTCATTCATGAGAGGCGCCCTGAGCGCTGAGAGGCACTCTGCG 426
Db 101 TyrGlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnGlyThrProGly 120
OY 427 TACAGGCCCCAGAGATCAGGCTGCTGCTGATGATGATGATGATGATGATGATGATGCTCC 486
Db 121 TyrGlnAlaProGlnIleArgProArgIleValTyrAspGlnLysValAspMetPheSer 140
OY 487 TATGGAATGGGCTGTAGAGTGTGCTGCAGAGACGCGCTGCACATGCGGCGACACACAG 546
Db 141 TyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHisIleGln 160
OY 547 CTCAGATTCGCAAGACCTGTCAGAGGCAATCCGCGCTTCTGCGGCGACCGAGAGAA 606
Db 161 LeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnGlnProGluLys 180
OY 607 GTGCAGTTCGCGGCGACTCGAGCGCTCATGATGAGAGTCTGGAGACTTAAGCCAGAGAG 666
Db 181 ValGlnPheArgArgLeuGlnAlaLeuMetMetGlnCysTrpAspThrLysProGluLys 200
OY 667 CGACCGTGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
Db 201 ArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThrPheMet 220
OY 727 TATGACTGTGCTGTGGAGAGAGACACCTTCTTCATCCAGGCGCCAGAGATACACC 786
Db 221 TyrGlnLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlnGlnLysThr 240
OY 787 GTGGTGTGTTGGGATGGAAGAAGAGAGTCCAGAGACTACAGGCTGTAACACAGAGAG 846
Db 241 ValValPheTrpAspGlyLysGlnLysSerArgAsnTyrThrValValAsnThrGlnLys 260
OY 847 GCGCCATGAGAGGTCGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
Db 261 GlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGlnLeuGln 280
OY 907 GTCCAGAGATCCTGTGAGACACCGACGAGACCGAGAGACGAGAAATCATCATATACACCTCAG 966

```

```

RESULT 3
PCT-US01-01435-82
; Sequence 82, Application PC/TUS0101435
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 21 human secreted proteins
; FILE REFERENCE: PS726PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01435
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/226,282
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 82

```

Db 281 ValGlnArgSerLeuThrPheAlaThrGluAspGlnIleuTyrlleuThrLeuLys 300
QY 967 GGCATGTGGCCCTTAACACACCCACAGGCGCTGGATACCCACCTGTCGACCTCG 1026
Db 301 GlyMetCysProLeuAsnThrProGlnAlaLeuAspThrProAlaValThrCys 320
QY 1027 TTTCTGGCCGCTGCTTATTAAGAAATTCCTACCTGCTTACGGGCGCTCGCCGAT 1086
Db 321 PheLeuAlaValProValIleuLysLysAsnSerTyrlleuValLeuAlaGlyLeuAlaAsp 340
QY 1087 GGGCTGTGGCTGTGTTCCTGGTGGCGGGACCCCAAGACAGCTGCTCTACCTG 1146
Db 341 GlyLeuValAlaValPheProValAlaArgGlyThrProLysAspSerCysSerTyrlleu 360
QY 1147 TGGTCACACACAGCCCAAGGCTCAAGTTCAGATCGCGGATGAACACAGCGGAGAC 1206
Db 361 CysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlnAspAlaArgGlnAsn 380
QY 1207 CCCTACCCAGTGAAAGCCATGAGAGTGTGCACACAGCGGCTGTAGGCTGTGACACAAAT 1266
Db 381 ProTyrlProValLysAlaMetGluValAlaAsnSerGlySerIuValItrPtyrSerAsn 400
QY 1267 GGGCCGGGCGCTCTGTTCATGACTGTGCTCCCTGAGATGTGCAGGGCGGTGAGCC 1326
Db 401 GlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeuGluPro 420
QY 1327 TACATGGCCCCCTCCATGGTACGTACGTGCTGACAGCTCTAGGGGACAGAGGGAGAG 1386
Db 421 TyrlMetAlaProSerMetValIthrSerValAlaCysSerSerIuGlyArgGlyGluGlu 440
QY 1387 GTGCTGTGGCTGTGATGACAGGCAAGCCACTCTGTGATGACACATCCACACCTAC 1446
Db 441 ValValItrPcysLeuAspAspLysAlaAsnSerLeuValMetItyrHisSerIthrItyr 460
QY 1447 CAGCTGTGTGCGCGGATCTGTGCGGGGTCCAGCCCTCGAGGACATGTTTCCCGTG 1506
Db 461 GlnLeuCysAlaArgTyrlPheCysGlyValProSerProLeuArgAspMetPheProVal 480
QY 1507 CGGCGCTGTGACACAGAACCCCGGACGACGACACAGGCGCAACCCAAAGGTGCTGAG 1566
Db 481 ArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysValProGlu 500
QY 1567 GGGGACATCCATCGCGAGCTGAGCATCATGATGATGAGAGCTGGGACGACAGATCTGT 1626
Db 501 GlyAspSerIleAlaAspValSerIleMetItyrSerGluIuLeuGlyIthrGlnIleLeu 520
QY 1627 ATCCACACAGAAATCACTACTGACTACTGCTCAATGCTCTACTCTCATCCACCC 1686
Db 521 IleHisGlnIuSerLeuThrAspTyrlCysSerMetSerItyrSerSerSerProPro 540
QY 1687 CGCGAGGCTGCGCAGAGTCCCTCAAGCTCCCAAGCTCCCAAGCAAGTTTCCAGTGTG 1746
Db 541 ArgGlnAlaIleAlaArgSerProSerSerLeuProSerSerProAlaSerSerSerVal 560
QY 1747 CCTTTCCCAACGAGCTGCGAGGAGCTCAGACATGTTACATAGCGCGGCTGTGCTCGAC 1806
Db 561 ProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAlaSerAsp 580
QY 1807 AGGCTGACAGATGACCTGACCCCATGAGCGGAGAGCTTACAGCCAGCACTGTCAGACC 1866
Db 581 ArgSerIuHisAspLeuThrProMetAspGlyIuThrPheSerGlnHisLeuGlnAla 600
QY 1867 GTGAAGATCTCGCGCTCAGAGACCTCATTTGGGTCCCAAGCGCGGTGAGATGTTATC 1926
Db 601 ValLysIleLeuAlaValAlaArgAspLeuIleTyrlValProArgArgGlyGlyAspValIle 620
QY 1927 GTCAATGGCCCTGGAAGAGATTCGAAGCCCAAGCGGCGGAGCATTCGCGCTTAAA 1986
Db 621 ValIleGlyLeuGluLysAspSerGlnAlaGlnArgGlyValIleAlaValLeuLys 640
QY 1987 GCCGAGAGCTGATCGCATGGGCTGTGATGATGATGATGATGATGATGATGATGATGAT 2046
Db 641 AlaArgGluLeuThrProHisGlyValLeuValAlaAspAlaValAlaValAspThr 660

QY 2047 GTGTGTGACACCTTTGAATAATGAAAAACAGAGATGTGCTGCGGCTGTGAGGGGCTGG 2106
Db 661 ValValCysThrPheGlnAsnGlnAsnThrGluItrPcysLeuAlaValItrArgGlyTrp 680
QY 2107 GCGCGCAGGAGCTTGACATTTTCTACAGTCTCTACGAGAGCTGGGCGGCTGAGGCT 2166
Db 681 GlyAlaArgGluPheAspIlePheTyrlGlnSerTyrlGluIuLeuGlyArgLeuGlnAla 700
QY 2167 TGCACATCGCAAGAGAAG 2184
Db 701 CysThrArgLysArgArg 706
RESULT 4
PCT-US01-11988-1620
Sequence 1620, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PFS46PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1620
LENGTH: 706
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-1620
Alignment Scores:
Pred. No.: 8,79e-223 Length: 706
Score: 3707.00 Matches: 706
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 91.35% Indels: 0
DB: 1 Gaps: 0
US-09-836-392-8_Copy_22_2205 (1-2184) x PCT-US01-11988-1620 (1-706)
QY 67 ATGCTGACAGCGCTGACAGCAACCCCTGCATGCTGAGCGCTCATCGGATCAGATCCACCCG 126
Db 1 MetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIleHisPro 20
QY 127 CTGTGCTGCGCCCTGAGAGCTGCGGCGGCTCAGAGCTCAACACCGTGTGCTCGGAGAC 186
Db 21 LeuCyrlPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSerGluAsn 40
QY 187 GCGAGAGATCTCTCTTATACCCCTGGAGACATGCTGCAACCAAAAAATAGCTACAG 246
Db 41 AlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnIleAlaTyrlGln 60
QY 247 ATGCGCTGCGGCGCTGAGCTACGTGCACAGAAAAACATCATCTTGTGTGACTGAACTG 306
Db 61 IleAlaSerGlyLeuMetAlaTyrlLeuHisLysLysAlaAsnIleIlePheCysAspLeuLysSer 80
QY 307 GACAACTTGTGTGTGTGTGCTCTTACAGTGAAGAGACATCAACATCAAGCTATCTGAC 366
Db 81 AspAsnIleLeuValItrPheSerLeuAspValLysGlnHisIleAsnIleLysLeuSerAsp 100
QY 367 TACGAGATTCGAGGCGAGTATTCATGAGAGCGGCGCTTAGCGGTGAGGCGACACTCTGCG 426
Db 101 TyrlGlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnGlyIthrProGly 120
QY 427 TACAGGCGCCAGAGATCAGGCGCTGCATGTATATGATGAGAAGATGATGATGATGATGAT 486
Db 121 TyrlGlnAlaProGluIleArgProArgIleValTyrlAspGlnLysValAspMetPheSer 140

```

OY 487 TATGAATGTCGCTCTACGAGTTCCTGTCAGAGCCGCTGACATGGGCAACACAG 546
    |||
Db 141 TyG1uMetValLeuTyG1uLeuLeuSerG1yGlnArgProAlaLeuG1yHISHisGln 160
OY 547 CTCAGATTCCAGAAAGCTGTCCAAAGGATCCGCGGTTCTGGGGCAGCGAGGAA 606
    |||
Db 161 LeuGlnLeuAlaLysLysLeuSerLysG1yL1eArgProValLeuG1yGlnProG1uG1u 180
OY 607 GTGCACTTCGGGAGCTCAGAGCGGCTATATGATGAGTGTGGGCACTAAAGCAGAG 666
    |||
Db 181 ValGlnPheArgLeuGlnAlaLeuMetLeuG1yStrPaspThrLysProG1uLys 200
OY 667 CGACCGCTGCGCTGTGCTGTGAGCAGATGAAGACCGGATTTGGCACTTCATG 726
    |||
Db 201 ArgProLeuAlaLeuSerValValSerG1uMetLysProThrPheAlaThrPheMet 220
OY 727 TATGAATGTCGCTGTGGAAGCAGACAGCCTTCTTCATCCAGGGCAGAGTACAC 786
    |||
Db 221 TyG1uLeuLysCysG1yGlnThrAlaPhePheSerSerGlnG1yGlnL1yThr 240
OY 787 GTGCTGTTGGATGGAAGAGAGATCCAGAACTACAGGCTGTGAACACAGAGAG 846
    |||
Db 241 ValValPheThrPaspG1yLysG1uGlnSerArgAsnTyThrValValAsnThrG1uLys 260
OY 847 GGCCTCATGAGGTGACAGAGATGCTGCTGCTGATGAAGTGAAGTGCACCTGCAG 906
    |||
Db 261 G1yLeuMetL1uValGlnArgMetCysCysProL1yMetLysValSerCysGlnLeuGln 280
OY 907 GTCCAGAGATCCCTGTGAGACAGCCAGGAGACAGAAATCTACATCTACACCTCAG 966
    |||
Db 281 ValGlnArgSerLeuThrPheAlaThrG1uAspL1uLysL1yL1yThrLeuLys 300
OY 967 GGCATGGCCCTTAAACACACCCAGCCCTGATCTCAGCTGTGCTGCTGCTGCTG 1026
    |||
Db 301 G1yMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAlaThrCys 320
OY 1027 TTCTTGGCCCTGCTGTATTAAAGAAATCTCCTAGGCTGTAGGGGCTGCGCAT 1086
    |||
Db 321 PheLeuAlaValProValL1eLysLysAsnSerTyLeuValLeuAlaG1yLeuAlaAsp 340
OY 1087 GGCCTGTGCTGTGTTCCGCTGCTGCGGAGCCCAAGAGACCTGCTCTACTG 1146
    |||
Db 341 G1yLeuValAlaValPheProValValArgG1yThrProLysAspSerCysSerTyLeu 360
OY 1147 TGTCAACACAGCCACAGCTCCAACTTCAAGCATCCGGATGAAGACGACGCGCAAC 1206
    |||
Db 361 CysSerHisThrAlaAsnArgSerLysPheSerL1eAlaAspG1uAspAlaArgGlnAsn 380
OY 1207 CCTACCCAGTGAAGGCGCATGAGGTGTCACAGCGGCTGTGAGTGTGATACGCAAT 1266
    |||
Db 381 ProTyProValLysAlaMetG1uValValAsnSerLysSerG1uValL1yP1ySerAsn 400
OY 1267 GGGCGGGGCTCTGTGATCGATGCTGCTCCCTGAGATCTGACAGCGGCTGAGAGCC 1326
    |||
Db 401 G1yProG1yLeuLeuValL1eAspCysAlaSerLeuG1uL1eCysArgArgLeuG1uPro 420
OY 1327 TACATGGCCCCCTCCATGCTTACGTAGTGTGAGCTGTGAGGCGACAGAGGAGAG 1386
    |||
Db 421 TyMetAlaProSerMetValL1ySerValValCysSerSerG1uG1yArgG1yG1uG1u 440
OY 1387 GTGCTGTGCTGATGATGAAGAGCCCACTCTGTGATGTATGCTACACTCCACACTAC 1446
    |||
Db 441 ValValL1yProCysLeuAspAspLysAlaAsnSerLeuValL1eC1yR1yHISSerThrThr 460
OY 1447 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
    |||
Db 461 GlnLeuLysAlaArgTyPheCysG1yValProSerProLeuArgAspMetPheProVal 480
OY 1507 CGGCGCTTGACAGAGAAACCCCGGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1566
    |||
Db 481 ArgProLeuAspThrG1uProProAlaL1eSerHisThrAlaAsnProLysValProG1u 500

```

```

OY 1567 GGGGACTCATCGCGGAGCGAGCATCATCATGATGAGAGAGCGGCGAGCAGATCCTG 1626
    |||
Db 501 G1yAspSerL1eAlaAspValSerL1eMetTySerL1uL1eG1yGlnG1uL1eLeu 520
OY 1627 ATCCACAGCAATCACTCACTGACTGACTGCTCCATGCTCTCTACTCTCATCCACACC 1686
    |||
Db 521 L1eHisGlnL1uSerLeuThrAspTyCysSerMetSerLysSerSerSerProPro 540
OY 1687 GCGCAGGCTCCAGAGTCCCTCCAGCCTCCAGCCTCCAGCCTCCAGCAGAGTTCCTGAGTGTG 1746
    |||
Db 541 ArgGlnAlaL1aArgSerProSerSerLeuProSerSerProAlaSerSerSerVal 560
OY 1747 CTTTTCACACCGAGCTGCGAGGAGCTCAGCATCTCATATCGCCGCTGCTGCTGCTG 1806
    |||
Db 561 ProPheSerThrAspLysG1uAspSerAspMetLeuHisThrProG1yAlaL1aSerAsp 580
OY 1807 AGGTCTGAGCATGACCTGACCCCATGAGCGGGAGAGACTTCAAGCCAGCAGCAGCAGGCC 1866
    |||
Db 581 ArgSerG1uHisAspLeuThrProMetAspG1yGlnThrPheSerGlnHisLeuGlnAla 600
OY 1867 GTGAAGATCTCGCCGCTGACAGACCTCATTTGGGTCCCGAGCGCGGTGAGATGTTATC 1926
    |||
Db 601 ValLysL1eLeuAlaValArgAspLeuL1eTyPValProArgArgG1yG1yAspValL1e 620
OY 1927 GTCATTCGCTGAGAGAGATTCGAGAGCCAGCGGGCGGAGTGCATTCGCTTAAAG 1986
    |||
Db 621 ValL1eG1yLeuG1uLysAspSerG1uAlaGlnArgG1yArgValL1eAlaValLeuLys 640
OY 1987 GCGCGAGCTGACTCCGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2046
    |||
Db 641 AlaArgG1uLeuThrProHisG1yValLeuValAspAlaAlaValAlaLysAspThr 660
OY 2047 GTTGTGTGACCTTTGAAATGAAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2106
    |||
Db 661 ValValCysThrPheG1uAsnG1uAsnThrG1uTrpCysLeuAlaValL1yPArgG1yTrp 680
OY 2107 GCGCGCAGGAGTTCGATTTCTTACAGTCTTACAGAGAGGAGGCGGCGCTGAGAGCT 2166
    |||
Db 681 G1yAlaArgL1uPheAspL1ePheTyGlnSerTyGlnL1eG1yArgLeuG1uAla 700
OY 2167 TGCACCTGCAAGAGAGG 2184
    |||
Db 701 CysThrArgLysArgArg 706
    |||

RESULT 5
US-09-833-245-1620
; Sequence 1620, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1620
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-1620

Alignment Scores:
Pred. No.: 8,79e-223
Score: 3707.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 91.35%

Length: 706
Matches: 706
Conservative: 0
Mismatch: 0
Indels: 0

```

Db 199 FPGIDPLEELFRKE--RPRRLPLPNCSEE-----LKDLIKKCLNKDPEKRPFAKEIL 251
QY 231 S 231
Db 252 N 252

RESULT 15

US-10-087-684-99
Sequence 99, Application US/10087684

GENERAL INFORMATION:
APPLICANT: Edinger, Shlomit R.
APPLICANT: MacDougall, John R.
APPLICANT: Millet, Isabelle
APPLICANT: Ellerman, Karen
APPLICANT: Stone, David J.
APPLICANT: Grose, William M.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
APPLICANT: Casman, Stacie, J.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Boldog, Ferenc L.
APPLICANT: Li, Li
APPLICANT: Padigaru, Muralidhara
APPLICANT: Mishra, Vishnu
APPLICANT: Shenoy, Suresh G.
APPLICANT: Rastelli, Luca
APPLICANT: Tchernev, Velizar T.
APPLICANT: Vernet, Corine A.M.
APPLICANT: Zerhusen, Bryan D.
APPLICANT: Malyankar, Urfel M.
APPLICANT: Guo, Xiaojia
APPLICANT: Miller, Charles E.
APPLICANT: Gangoli, Esha A.
TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-214 CIP
CURRENT APPLICATION NUMBER: US/10/087,684
CURRENT FILING DATE: 2003-03-10
PRIOR APPLICATION NUMBER: 60/253,834
PRIOR FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 60/250,926
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 60/264,180
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/313,656
PRIOR FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/327,456
PRIOR FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 220
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 99
LENGTH: 256
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Domain
US-10-087-684-99

Query Match 7.1%; Score 271; DB 6; Length 256;
Best Local Similarity 32.4%; Pred. No. 3.5e-13;
Matches 78; Conservative 42; Mismatches 93; Indels 28; Gaps 7;

QY 2 LRHLBATDAKNSFROEASMLHQPCTVALIGI--SIHPLCFALDELAPLSSNTVL 59
Db 29 IKILKRSJSEKKRFLRLSHPNIVRLGVFEEDDLTYLWMEYEGGDLFDYL 88
QY 60 SENARDSFPIGLMLOTOKIAYOIASGLAYIHKKNITFCDLKSDNLTWSLDYKEHINIK 119
Db 89 RRNG-----LLSEKEKAKIALQILRGLEYLHSGIYVHRDLKPENIL-----LDENGTVK 138

QY 120 LSDYGISRQ---SFHEGALVGEPTPGYQAPRIRPRIYDEKVDMSFGMVLIELSCORP 176
Db 139 IADEGLARKLESSSYEXLITFVGTPETWAPPEYLEGSGYSKVDVMSIGVILYELTGKLP 198
QY 177 ALGHOLQIAKKLSKGIKRPVYGOP-----BEVGFRLQALMECMPTKPKRPLALSYV 230
Db 199 FPGIDPLEELFRKE--RPRRLPLPNCSEE-----LKDLIKKCLNKDPEKRPFAKEIL 251
QY 231 S 231
Db 252 N 252

Search completed: April 8, 2003, 10:33:52
Job time : 86 secs

Pred. NO. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

15222.671 Million cell updates/sec

Sequence: 1 atgctgagcacctgcggc.....ctgcactcgcaagagaag 2184

BLOSUM62			
Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
delop	6.0	delext	7.0

Total number of hits satisfying chosen parameters: 9138288

Post-processing: Minimum Match 08

Listing first 45 summaries

Command line parameters:
-MODEL firstnot 230 model1

```

model_loader=12p_model_DEVEX11n
-O=/cgn2.1/USPO.spool/USO9836392/runat.08042003_090329_22865/app_query.fasta.1.23759
-DB=reading_patents_Aa_Main -QFMT=fastan -SUFFIX=12p_rapm -MINMATCH=0.1
-LOOBL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRIX=bisnumm62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPOST=500 -MINLEN=0
-MAXLEN=2000000000 -USER=0509836392 -CGEN.1.1.255 -runat.08042003_090329_22865
-NCPU=6 -ICRU=3 -NO_XIPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database :

- 1: /cgn2_6/prodata/1/paa/PCRTUS.COMB.pcp.*
- 2: /cgn2_6/prodata/1/paa/US05.COMB.pcp.*
- 3: /cgn2_6/prodata/1/paa/US07.COMB.pcp.*
- 4: /cgn2_6/prodata/1/paa/US08_01.COMB.pcp.*
- 5: /cgn2_6/prodata/1/paa/US08_1.COMB.pcp.*
- 6: /cgn2_6/prodata/1/paa/US08_2.COMB.pcp.*
- 7: /cgn2_6/prodata/1/paa/US08_3.COMB.pcp.*
- 8: /cgn2_6/prodata/1/paa/US08_4.COMB.pcp.*
- 9: /cgn2_6/prodata/1/paa/US08_5.COMB.pcp.*
- 10: /cgn2_6/prodata/1/paa/US08_6.COMB.pcp.*
- 11: /cgn2_6/prodata/1/paa/US08_7.COMB.pcp.*
- 12: /cgn2_6/prodata/1/paa/US08_8.COMB.pcp.*
- 13: /cgn2_6/prodata/1/paa/US09.COMB.pcp.*
- 14: /cgn2_6/prodata/1/paa/US09_01.COMB.pcp.*
- 15: /cgn2_6/prodata/1/paa/US09_1.COMB.pcp.*
- 16: /cgn2_6/prodata/1/paa/US09_2.COMB.pcp.*
- 17: /cgn2_6/prodata/1/paa/US09_3.COMB.pcp.*
- 18: /cgn2_6/prodata/1/paa/US09_4.COMB.pcp.*
- 19: /cgn2_6/prodata/1/paa/US09_5.COMB.pcp.*
- 20: /cgn2_6/prodata/1/paa/US09_6.COMB.pcp.*
- 21: /cgn2_6/prodata/1/paa/US09_7.COMB.pcp.*
- 22: /cgn2_6/prodata/1/paa/US09_8.COMB.pcp.*
- 23: /cgn2_6/prodata/1/paa/US10.COMB.pcp.*
- 24: /cgn2_6/prodata/1/paa/US10_01.COMB.pcp.*
- 25: /cgn2_6/prodata/1/paa/US10_2.COMB.pcp.*
- 26: /cgn2_6/prodata/1/paa/US10_3.COMB.pcp.*
- 27: /cgn2_6/prodata/1/paa/US60.COMB.pcp.*

Pred. NO. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
---------------	-------	----------------	--------	----	----	-------------

1	3822	94.2	728	1	PCT-US00-28066-21	Sequence 21, Appl.
2	3822	94.2	728	22	US-09-836-192-21	Sequence 21, Appl.
3	3707	91.4	706	1	PCT-US01-01433-82	Sequence 82, Appl
4	3707	91.4	706	1	PCT-US01-11988-1620	Sequence 1620, Ap
5	3707	91.4	706	22	US-09-833-445-1620	Sequence 1620, Ap
6	3700	91.2	706	1	PCT-US01-01335-59	Sequence 59, Appl
7	3700	91.2	706	1	PCT-US01-11988-1621	Sequence 1621, Ap
8	3700	91.2	706	22	US-09-833-445-1621	Sequence 1621, Ap
9	3574.5	88.1	1987	25	US-10-132-882-6	Sequence 6, Appl
10	3574.5	88.1	2013	25	US-10-132-882-2	Sequence 2, Appl
11	3574.5	88.1	2014	25	US-10-132-882-4	Sequence 8, Appl
12	3574.5	88.1	2040	25	US-10-132-882-8	Sequence 4, Appl
13	3349.5	82.5	809	1	PCT-US01-05039-38410	Sequence 38410, A
14	1769	41.4	320	1	PCT-US02-00510-98-92	Sequence 982, App
15	435	10.7	1050	27	US-60-155-811-774	Sequence 774, App
16	383.5	9.5	288	27	US-60-142-844-557	Sequence 557, App
17	380.5	9.4	622	27	US-60-145-137-532	Sequence 532, App
18	364.5	9.0	2308	20	US-09-614-150-11484	Sequence 11484, A
19	364.5	9.0	2308	27	US-60-173-464-9298	Sequence 9299, App
20	364.5	9.0	2308	27	US-60-191-637-11516	Sequence 11516, A
21	364.5	9.0	2308	27	US-60-191-681-9008	Sequence 9008, App
22	364.5	9.0	2321	27	US-60-167-217-11502	Sequence 11502, A
23	356.5	8.8	266	27	US-60-146-315-7749	Sequence 749, App
24	348.5	8.6	826	27	US-60-213-178-794	Sequence 794, App
25	338	8.3	848	27	US-60-213-178-831	Sequence 831, App
26	337.5	8.3	847	21	US-09-791-537-391	Sequence 391, App
27	337.5	8.3	847	25	US-10-143-133-2	Sequence 2, Appl
28	337.5	8.3	847	27	US-60-350-061-241	Sequence 241, Appl
29	335.5	8.3	915	1	PCT-US02-10522-50	Sequence 48, Appl
30	335.5	8.3	915	25	US-10-135-482-48	Sequence 48, Appl
31	327.5	8.1	183	27	US-60-139-671-197	Sequence 197, App
32	327.5	8.1	183	27	US-60-140-806-1118	Sequence 1118, App
33	327.5	8.1	656	27	US-60-215-651-1	Sequence 1, Appl
34	327.5	8.1	911	1	PCT-US02-10522-50	Sequence 50, Appl
35	327.5	8.1	911	25	US-10-115-482-50	Sequence 50, Appl
36	327.5	8.1	1818	27	US-60-345-773-2	Sequence 2, Appl
37	327.5	8.1	1824	27	US-60-345-773-2	Sequence 5, Appl
38	327.5	8.1	2527	1	PCT-US02-24521-6	Sequence 6, Appl
39	320	8.0	65	27	US-60-196-711-2337	Sequence 233, App
40	320	7.9	67	27	US-60-182-074-263	Sequence 263, App
41	317.5	7.8	953	21	US-09-791-537-917	Sequence 917, App
42	317.5	7.8	953	24	US-10-042-865-98	Sequence 98, Appl
43	316.5	7.8	844	1	PCT-US01-14827-12358	Sequence 12358, A
44	315.5	7.8	1097	27	US-60-234-802-6	Sequence 6, Appl
45	315	7.8	369	19	US-09-513-9966-24999	Sequence 24999, A

ALIGNMENTS

```

RESULT 1
PCT-US00-28066-21
: Sequence 21, Application PC/TUS0028066
:
: GENERAL INFORMATION:
:   APPLICANT: Human Genome Sciences, Inc.
:   TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
:   TITLE OF INVENTION: Antibodies
:   FILE REFERENCE: PT020PCF
:   CURRENT APPLICATION NUMBER: PCT/US00/28066
:   CURRENT FILING DATE: 2000-10-12
:   PRIOR APPLICATION NUMBER: 60/159,542
:   PRIOR FILING DATE: 1999-10-15
:   PRIOR APPLICATION NUMBER: 60/165,914
:   PRIOR FILING DATE: 1999-11-17
:   PRIOR APPLICATION NUMBER: 60/189,027

```

PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 728
TYPE: PRT
ORGANISM: Homo sapiens
PCT-0500-28066-21

Alignment Scores:
Pred. No.: 5,83e-230 Length: 728
Score: 3822.00 Matches: 728
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.18% Indels: 0
DB: Gaps: 0

US-09-836-392-8_COPY.22_2205 (1-2184) x PCT-US00-28066-21 (1-728)

QY 1 ATGTGAGGACCTGGCGGCGCAGCATGAGAACTTCCTCGAGTTCGGCAGAG 60
Db 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheserGluPheArgInGlu 20
QY 61 GCCAGCATGCTGCAGCGGCTGACAGACCCCTGCATCTGGCGCTATCGCATCGATC 120
Db 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
QY 121 CACCGGCTCTGCTCGCCCTGAGAGCTCGCGGCTCAGCAGCCTCAGACCGGTGCTGC 180
Db 41 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerIleAsnThrValLeuSer 60
QY 181 GAGAACGCCAGAGATTCCTCTTATACCCCTGGGACACATGCTCACCCAAAATAACC 240
Db 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 80
QY 241 TACGATGCGCTGGCGGCTGCGCTACCTGCACAAAGAAAACATCATCTCTGTGACCTG 300
Db 81 TyrGlnIleAlaSerGlyLeuAlaIleTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
QY 301 AAGTCGGAACATCTTGTTGGTGTGCTCCCTTACAGCTCAAGAGACACATCAACATCA 360
Db 101 LysSerAspAsnIleLeuValIlePheSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
QY 361 TCTGACTACGGGATTTGAGCAGCATTCATTCATAGAGCGCCCTAGGCTGGAGGGCACT 420
Db 121 SerAspTyrGlyIleSerArgIleSerPheHisGlnGlyAlaLeuGlyValGlnGlyThr 140
QY 421 CCTGCTACGAGCGGCCGAGATCAGCGCTGCGCATGTATATGATGAGAAGTATATG 480
Db 141 ProGlyTyrGlnAlaProGlnIleArgProAlaIleValIleTyrAspGlnLysValAspMet 160
QY 481 TTCTCTATGGAATGTGCTCTACAGATTGCTGTCAGAGACAGCGGCTGACATGGGCGAC 540
Db 161 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnLys 180
QY 541 CACCACTCCAGATTGCCAAGAGCTGTCAAGGGCATCCCGGCTTCTGGGGCAGCGG 600
Db 181 HisGlnLeuGlnIleAlaLysIleLysLeuSerLysGlyIleArgProValLeuGlnPro 200
QY 601 GAGGAAGTCAAGTTCGGCGGCTGAGCGGCTCATGATGATGAGTGGGAGACAATAAGCA 660
Db 201 GluGlnValGlnPheArgIleGlnLeuGlnAlaLeuMetGlyCysTyrAspThrLysPro 220
QY 661 GAGAGACGACCGCTGGCGCTGCTGGTGTGAGCCAGATGAGAGACCCGACATTTTGGCAC 720
Db 221 GluLysArgProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThr 240
QY 721 TTCATGTATGACTGTGTGTGGAGACAGACAGCCTTCTCTCATCCAGCGCCAGAG 780
Db 241 PheMetTyrGlnLeuCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGln 260
QY 781 TACACCGGTGTGTGGATGAGAAAGAGAGTCCAGAACTACAGGCTGTGAGACCA 840

Db 261 TyrThrValValPheTyrAspGlyLysGlnGlnSerArgAsnTyrThrValValAsnThr 280
QY 841 GAGAAAGGCGCTCATGAGAGTGCAGAGATGTGCTCCCTGGAGTGAAGTGTGCGCAG 900
Db 281 GluLysGlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
QY 901 CTCAGGTCAGAGATCCCTGTGAGACGCCAGCAGAGACGAAATCTACATCTACACC 960
Db 301 LeuGlnValGlnArgSerLeuTyrThrAlaThrGlnAspGlnLysIleTyrIleTyrThr 320
QY 961 CTCAGGCGCATGTGCCCTTAAACACACCCCAAGGCTTGGATCTCAGCTGTGCTC 1020
Db 321 LeuLysGlyMetCysProLeuAsnThrProGlnAlaLeuAspThrProAlaValVal 340
QY 1021 ACCTGCTCTTGGCGGCTGCTGTATTAAGAAATCTACCTGCTTACCGGCGCTC 1080
Db 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
QY 1081 GCCGATGGGCTTGTGCTGTGTTCCCGTGTGCGGCGCACCACCAAGACAGCTGCTCC 1140
Db 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
QY 1141 TACCTGTCTACACACAGCCACAGCTCCAGTTCAGCATCCGGGATGAAGACGACGG 1200
Db 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
QY 1201 CAGAACCCCTACCCAGTGAAGGCGCATGAGGTGTGTCAACACGCGCTGTAGAGTGTG 1260
Db 401 GlnAsnProTyrProValLysAlaMetGlnValValAsnSerGlySerGlnValIleTyr 420
QY 1261 AGCAATGGGCGGCGCTCTTGTGATGACTGTGCTCCTCGAGATCTGCAGCGGCTG 1320
Db 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgArgLeu 440
QY 1321 GAGCCCTACATGAGCGCCCTCCATGTGTACGTCATCTGTGCTGAGCGGCGCAGGG 1380
Db 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgGly 460
QY 1381 GAGAGGTGCTGTGCTGTGATGAGACAGGCGCAACTCTGTGTGATGTACCATCCACC 1440
Db 461 GluGlnValValIleProCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
QY 1441 ACCTACCAAGCTGTGTGCGGCTACTTCTGCGGGGTCCCGACCCCTCAGGAGATGTTT 1500
Db 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
QY 1501 CCGGTGGGCGCTTGGAGACAGGACCCCGGACCGACGACGAGGCAACCCAAAGTG 1560
Db 501 ProValAlaProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
QY 1561 CCTGAGGGGAGCTCCATCGCGAGAGTGAAGCATGTATGACATGAGAGAGTGGGACGAC 1620
Db 521 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGlnLeuGlyThrGln 540
QY 1621 ATCTGATCCACAGGAATCACTGACTACTGTGTCATGTCTCTACTCTCTATCC 1680
Db 541 IleLeuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerTyrSerSer 560
QY 1681 CCACCGGCGAGGCTGCGAGGCTGCCCTCAAGCTCCCAAGCTCCCAAGAAATTTCC 1740
Db 561 ProProAlaGlnAlaAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSer 580
QY 1741 AGTGTGCTTCTTCCACCGACTCGAGAGACTCAGACATGTATACATGACCGCGGTGCTGC 1800
Db 581 SerValProPheSerThrAspCysGluAspSerSerMetLeuHisThrProGlyAlaAla 600
QY 1801 TCGGACAGGTGTGAGCATGACCTGACCCCAATGAGCGGGAGACCTTACGACACACTG 1860
Db 601 SerAspArgSerGlnHisAspLeuThrProMetAspArgGlnThrPheSerGlnHisLeu 620
QY 1861 CAGGCGTGAAGATCTCGCGGCTGAGAGACCTCATTTGGGCTCCGACGCGGTGAGAT 1920
Db 621 GlnAlaValLysIleLeuAlaValAlaArgAspLeuIleThrProAlaArgGlyLysPhe 640

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using SW model

Run on: April 8, 2003, 10:23:14 ; Search time 322 Seconds

(without alignments)
1457.657 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822
Sequence: 1 MRLHRTADKMFSEFROE.....IFYQSYELGRLKTRRR 728

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main.*
1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	3822	100.0	728	1	PCT-US00-28066-21
2	3822	100.0	728	22	US-09-836-392-21
3	3712	97.1	706	1	PCT-US01-01435-82
4	3712	97.1	706	1	PCT-US01-11988-1620
5	3712	97.1	706	22	US-09-833-245-1620
6	3705	96.9	706	1	PCT-US01-01435-59

7	3705	96.9	706	1	PCT-US01-11988-1621	Sequence 1621, Ap
8	3705	96.9	706	22	US-09-833-245-1621	Sequence 1621, Ap
9	3574.5	93.5	1987	25	US-10-132-382-6	Sequence 6, Appl
10	3574.5	93.5	2013	25	US-10-132-382-2	Sequence 2, Appl
11	3574.5	93.5	2014	25	US-10-132-382-8	Sequence 8, Appl
12	3574.5	93.5	2040	25	US-10-132-382-4	Sequence 4, Appl
13	2997	78.4	809	1	PCT-US01-08631-38410	Sequence 38410, A
14	1679	43.9	320	1	PCT-US02-05109-982	Sequence 982, App
15	435	11.4	1050	27	US-60-155-811-774	Sequence 774, App
16	383.5	10.0	288	27	US-60-142-844-557	Sequence 557, App
17	380.5	10.0	622	27	US-60-145-157-532	Sequence 532, App
18	364.5	9.5	2308	20	US-09-614-150-11464	Sequence 11464, A
19	364.5	9.5	2308	27	US-60-173-464-9298	Sequence 9298, Ap
20	364.5	9.5	2308	27	US-60-191-637-11516	Sequence 11516, A
21	364.5	9.5	2308	27	US-60-191-681-9008	Sequence 9008, Ap
22	364.5	9.5	2321	27	US-60-167-217-11502	Sequence 11502, A
23	356.5	9.3	266	27	US-60-146-315-749	Sequence 749, App
24	335.5	8.8	915	1	PCT-US02-10522-48	Sequence 48, Appl
25	335.5	8.8	915	25	US-10-115-482-48	Sequence 48, Appl
26	327.5	8.6	183	27	US-60-139-671-197	Sequence 197, App
27	327.5	8.6	183	27	US-60-140-806-1118	Sequence 1118, App
28	327.5	8.6	656	27	US-60-215-651-1	Sequence 1, Appl
29	327.5	8.6	911	25	PCT-US02-10522-50	Sequence 50, Appl
30	327.5	8.6	911	25	US-10-115-482-50	Sequence 50, Appl
31	327.5	8.6	1818	27	US-60-345-773-5	Sequence 2, Appl
32	327.5	8.6	1824	27	US-60-345-773-5	Sequence 5, Appl
33	327.5	8.6	2527	1	PCT-US02-24521-6	Sequence 2337, Ap
34	320	8.5	65	27	US-60-196-711-2337	Sequence 263, App
35	320	8.4	844	1	PCT-US01-14827-12358	Sequence 12358, A
36	316.5	8.3	869	19	US-09-513-966A-2499	Sequence 2499, A
37	315	8.2	369	20	US-09-649-866A-798	Sequence 798, App
38	315	8.2	374	19	US-09-513-966A-2498	Sequence 2498, A
39	315	8.2	374	20	US-09-649-866A-797	Sequence 797, App
40	315	8.2	412	19	US-09-513-966A-2497	Sequence 2497, A
41	315	8.2	412	19	US-09-573-655A-1028	Sequence 1028, Ap
42	315	8.2	412	19	US-09-573-655B-1028	Sequence 1028, Ap
43	315	8.2	412	20	US-09-649-866A-796	Sequence 796, App
44	315	8.2	412	20	US-60-219-150-8	Sequence 8, Appl
45	315	8.2	412	27		

ALIGNMENTS

RESULT 1
PCT-US00-28066-21
Sequence 21, Application PC/TUS0028066
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc..
TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptid
FILE REFERENCE: PRO20PCT
CURRENT APPLICATION NUMBER: PCT/US00/28066
CURRENT FILING DATE: 2000-10-12
PRIOR APPLICATION NUMBER: 60/159,542
PRIOR FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: 60/165,914
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: 60/189,027
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 728
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-28066-21

Query Match 100.0% Score 3822; DB 1; Length 728;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRLHRTADKMFSEFROEASMLHALQPCIVALLIGISIHPLCFALPLASINTVLS 60

```

|||||
Db 1 MNRHLATAMKNFSEFEDEASMLHALQPCIALIGISIHPLCFALLETAPLSLNTVLS 60
OY 61 ENARSSFIPLGMLTOKIAYQIASGLATVLRHKKNIIFCOLKSDNIIWMSLDVKEHINIKL 120
Db 61 ENARSSFIPLGMLTOKIAYQIASGLATVLRHKKNIIFCOLKSDNIIWMSLDVKEHINIKL 120
OY 121 SDGISROSFHEGALGVEGTGPGYQAEIRPRIYDEKVMFSGVMVLYELLSGORPALGH 180
Db 121 SDGISROSFHEGALGVEGTGPGYQAEIRPRIYDEKVMFSGVMVLYELLSGORPALGH 180
OY 181 HOLOIAKKLSKIGIRPVLAGPPEEVOFRRLQALMMECHDTPEKRPPLALSVSQKDPTEFAT 240
Db 181 HOLOIAKKLSKIGIRPVLAGPPEEVOFRRLQALMMECHDTPEKRPPLALSVSQKDPTEFAT 240
OY 241 FMYELCCGKOTAFSSQGGETVYVFDGKEESRNTVVTTEKGLMEVORCCPGMKVSCQ 300
Db 241 FMYELCCGKOTAFSSQGGETVYVFDGKEESRNTVVTTEKGLMEVORCCPGMKVSCQ 300
OY 301 LOVORSLMTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLAVPVIKKNSYLVLAGL 360
Db 301 LOVORSLMTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLAVPVIKKNSYLVLAGL 360
OY 361 AGGLVAFVVGKTPDSCSYLCSHTANRKSSTADEARONPYVYKAMEVNSGSEWY 420
Db 361 AGGLVAFVVGKTPDSCSYLCSHTANRKSSTADEARONPYVYKAMEVNSGSEWY 420
OY 421 SNGPGLVLDCALEICRLEPEYMAPSMVTSVYCSSEGGEEVWCLDDKANSLVYHST 480
Db 421 SNGPGLVLDCALEICRLEPEYMAPSMVTSVYCSSEGGEEVWCLDDKANSLVYHST 480
OY 481 TYOLCARFYCGVPSPLRDMFPVPRPLDTEPPAASHANPVYPEGDSIADVSIMYSEELGQ 540
Db 481 TYOLCARFYCGVPSPLRDMFPVPRPLDTEPPAASHANPVYPEGDSIADVSIMYSEELGQ 540
OY 541 ILIHOSLTDYCSMSYSSPPROAARSPSLPSSPASSSVPEFSDCEDSDMLHTPGAA 600
Db 541 ILIHOSLTDYCSMSYSSPPROAARSPSLPSSPASSSVPEFSDCEDSDMLHTPGAA 600
OY 601 SDRSEHDLTPMDGETFSQHLQAVKILAVRDLIWPARRGGDVIVIGLEKSEARGVIAN 660
Db 601 SDRSEHDLTPMDGETFSQHLQAVKILAVRDLIWPARRGGDVIVIGLEKSEARGVIAN 660
OY 661 LKARELTPHGVLDAAVAVAKDVTVCFFENENTEMCLAVRGARGAREFDIFYOSYEBELGRL 720
Db 661 LKARELTPHGVLDAAVAVAKDVTVCFFENENTEMCLAVRGARGAREFDIFYOSYEBELGRL 720
OY 721 EACTRRRR 728
Db 721 EACTRRRR 728

RESULT 2
US-09-836-392-21
; Sequence 21, Application US/09836392
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides, a
; FILE REFERENCE: PTO20P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21

```

```

; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-392-21

Query Match 100.0%; Score 3822; DB 22; Length 728;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNRHLATAMKNFSEFEDEASMLHALQPCIALIGISIHPLCFALLETAPLSLNTVLS 60
Db 1 MNRHLATAMKNFSEFEDEASMLHALQPCIALIGISIHPLCFALLETAPLSLNTVLS 60
OY 61 ENARSSFIPLGMLTOKIAYQIASGLATVLRHKKNIIFCOLKSDNIIWMSLDVKEHINIKL 120
Db 61 ENARSSFIPLGMLTOKIAYQIASGLATVLRHKKNIIFCOLKSDNIIWMSLDVKEHINIKL 120
OY 121 SDGISROSFHEGALGVEGTGPGYQAEIRPRIYDEKVMFSGVMVLYELLSGORPALGH 180
Db 121 SDGISROSFHEGALGVEGTGPGYQAEIRPRIYDEKVMFSGVMVLYELLSGORPALGH 180
OY 181 HOLOIAKKLSKIGIRPVLAGPPEEVOFRRLQALMMECHDTPEKRPPLALSVSQKDPTEFAT 240
Db 181 HOLOIAKKLSKIGIRPVLAGPPEEVOFRRLQALMMECHDTPEKRPPLALSVSQKDPTEFAT 240
OY 241 FMYELCCGKOTAFSSQGGETVYVFDGKEESRNTVVTTEKGLMEVORCCPGMKVSCQ 300
Db 241 FMYELCCGKOTAFSSQGGETVYVFDGKEESRNTVVTTEKGLMEVORCCPGMKVSCQ 300
OY 301 LOVORSLMTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLAVPVIKKNSYLVLAGL 360
Db 301 LOVORSLMTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLAVPVIKKNSYLVLAGL 360
OY 361 AGGLVAFVVGKTPDSCSYLCSHTANRKSSTADEARONPYVYKAMEVNSGSEWY 420
Db 361 AGGLVAFVVGKTPDSCSYLCSHTANRKSSTADEARONPYVYKAMEVNSGSEWY 420
OY 421 SNGPGLVLDCALEICRLEPEYMAPSMVTSVYCSSEGGEEVWCLDDKANSLVYHST 480
Db 421 SNGPGLVLDCALEICRLEPEYMAPSMVTSVYCSSEGGEEVWCLDDKANSLVYHST 480
OY 481 TYOLCARFYCGVPSPLRDMFPVPRPLDTEPPAASHANPVYPEGDSIADVSIMYSEELGQ 540
Db 481 TYOLCARFYCGVPSPLRDMFPVPRPLDTEPPAASHANPVYPEGDSIADVSIMYSEELGQ 540
OY 541 ILIHOSLTDYCSMSYSSPPROAARSPSLPSSPASSSVPEFSDCEDSDMLHTPGAA 600
Db 541 ILIHOSLTDYCSMSYSSPPROAARSPSLPSSPASSSVPEFSDCEDSDMLHTPGAA 600
OY 601 SDRSEHDLTPMDGETFSQHLQAVKILAVRDLIWPARRGGDVIVIGLEKSEARGVIAN 660
Db 601 SDRSEHDLTPMDGETFSQHLQAVKILAVRDLIWPARRGGDVIVIGLEKSEARGVIAN 660
OY 661 LKARELTPHGVLDAAVAVAKDVTVCFFENENTEMCLAVRGARGAREFDIFYOSYEBELGRL 720
Db 661 LKARELTPHGVLDAAVAVAKDVTVCFFENENTEMCLAVRGARGAREFDIFYOSYEBELGRL 720
OY 721 EACTRRRR 728
Db 721 EACTRRRR 728

RESULT 3
PCT-US01-01435-82
; Sequence 82, Application PC/US0101435
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 21 human secreted proteins
; FILE REFERENCE: P5726PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01435
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/226,282
; PRIOR FILING DATE: 2000-08-18

```

NUMBER OF SEQ ID NOS: 145
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 82
LENGTH: 706
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-01435-82

Query Match 97.1%; Score 3712; DB 1; Length 706;
Best Local Similarity 100.0%; Pred. No. 1.1e-312;
Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

23 MHALQHPCIVALLIGISHPICFALELAPLSLNTVLSNADSSFIPLGHMLTKRIAYQ 82
1 MHALQHPCIVALLIGISHPICFALELAPLSLNTVLSNADSSFIPLGHMLTKRIAYQ 60
83 IASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKLSDVGISRSQSFEGALGVEGTPG 142
61 IASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKLSDVGISRSQSFEGALGVEGTPG 120
143 YQAPETIRPIYDEKVDNFSTGMVLYELISGORPALGHHQLOIAKLSKIRPVLGQPEE 202
121 YQAPETIRPIYDEKVDNFSTGMVLYELISGORPALGHHQLOIAKLSKIRPVLGQPEE 180
203 VQFRRLQALMBCMDTPKPKRPLALSVYSQMDPTFATFMYELCCGKQAFSSQGOEYT 262
181 VQFRRLQALMBCMDTPKPKRPLALSVYSQMDPTFATFMYELCCGKQAFSSQGOEYT 240
263 VVFMGKESRNYTYVNTKEGLMEVQRMCCPGMKVSCQLOVQRSMTATEDOKIITYTLK 322
241 VVFMGKESRNYTYVNTKEGLMEVQRMCCPGMKVSCQLOVQRSMTATEDOKIITYTLK 300
323 GMCPLNTPOQALDTPAVVTCFLAVPYIKKNSLYLAGLADGLVAVFPVVRGTPKDCSYL 382
301 GMCPLNTPOQALDTPAVVTCFLAVPYIKKNSLYLAGLADGLVAVFPVVRGTPKDCSYL 360
383 CSHTANRSKFSIADDEBARONPYPVKAMEVYNSGSEWYNSNGGLVITDCALEICRRLP 442
361 CSHTANRSKFSIADDEBARONPYPVKAMEVYNSGSEWYNSNGGLVITDCALEICRRLP 420
443 YVAPSMYTSVYVSSSEBGEVYVWCLDDKANSILVMYHSTTYQCARFCVPSPLRDMFV 502
421 YVAPSMYTSVYVSSSEBGEVYVWCLDDKANSILVMYHSTTYQCARFCVPSPLRDMFV 480
503 RPLDTPPAASHTANPKVPEGDSIADVSIYSEELGTQILIHQESLTDYCSMSYSSSP 562
481 RPLDTPPAASHTANPKVPEGDSIADVSIYSEELGTQILIHQESLTDYCSMSYSSSP 540
563 ROAARSPSLPSSPASSSSVPESTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQHOLA 622
541 ROAARSPSLPSSPASSSSVPESTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQHOLA 600
623 VVILAVRDLIWPVRRGVDIVIGLEKDSQGRVIAVAKARELPHGVLYDAVAAVAKDT 682
601 VVILAVRDLIWPVRRGVDIVIGLEKDSQGRVIAVAKARELPHGVLYDAVAAVAKDT 660
683 VVCTFENENTEMCLAVWGMGAREFDIFYOSEYELGRLDCTCRKR 728
661 VVCTFENENTEMCLAVWGMGAREFDIFYOSEYELGRLDCTCRKR 706

RESULT 4
PCT-US01-11988-1620
Sequence 1620, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931

PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 1620
LENGTH: 706
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-1620

Query Match 97.1%; Score 3712; DB 1; Length 706;
Best Local Similarity 100.0%; Pred. No. 1.1e-312;
Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

23 MHALQHPCIVALLIGISHPICFALELAPLSLNTVLSNADSSFIPLGHMLTKRIAYQ 82
1 MHALQHPCIVALLIGISHPICFALELAPLSLNTVLSNADSSFIPLGHMLTKRIAYQ 60
83 IASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKLSDVGISRSQSFEGALGVEGTPG 142
61 IASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKLSDVGISRSQSFEGALGVEGTPG 120
143 YQAPETIRPIYDEKVDNFSTGMVLYELISGORPALGHHQLOIAKLSKIRPVLGQPEE 202
121 YQAPETIRPIYDEKVDNFSTGMVLYELISGORPALGHHQLOIAKLSKIRPVLGQPEE 180
203 VQFRRLQALMBCMDTPKPKRPLALSVYSQMDPTFATFMYELCCGKQAFSSQGOEYT 262
181 VQFRRLQALMBCMDTPKPKRPLALSVYSQMDPTFATFMYELCCGKQAFSSQGOEYT 240
263 VVFMGKESRNYTYVNTKEGLMEVQRMCCPGMKVSCQLOVQRSMTATEDOKIITYTLK 322
241 VVFMGKESRNYTYVNTKEGLMEVQRMCCPGMKVSCQLOVQRSMTATEDOKIITYTLK 300
323 GMCPLNTPOQALDTPAVVTCFLAVPYIKKNSLYLAGLADGLVAVFPVVRGTPKDCSYL 382
301 GMCPLNTPOQALDTPAVVTCFLAVPYIKKNSLYLAGLADGLVAVFPVVRGTPKDCSYL 360
383 CSHTANRSKFSIADDEBARONPYPVKAMEVYNSGSEWYNSNGGLVITDCALEICRRLP 442
361 CSHTANRSKFSIADDEBARONPYPVKAMEVYNSGSEWYNSNGGLVITDCALEICRRLP 420
443 YVAPSMYTSVYVSSSEBGEVYVWCLDDKANSILVMYHSTTYQCARFCVPSPLRDMFV 502
421 YVAPSMYTSVYVSSSEBGEVYVWCLDDKANSILVMYHSTTYQCARFCVPSPLRDMFV 480
503 RPLDTPPAASHTANPKVPEGDSIADVSIYSEELGTQILIHQESLTDYCSMSYSSSP 562
481 RPLDTPPAASHTANPKVPEGDSIADVSIYSEELGTQILIHQESLTDYCSMSYSSSP 540
563 ROAARSPSLPSSPASSSSVPESTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQHOLA 622
541 ROAARSPSLPSSPASSSSVPESTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQHOLA 600
623 VVILAVRDLIWPVRRGVDIVIGLEKDSQGRVIAVAKARELPHGVLYDAVAAVAKDT 682
601 VVILAVRDLIWPVRRGVDIVIGLEKDSQGRVIAVAKARELPHGVLYDAVAAVAKDT 660
683 VVCTFENENTEMCLAVWGMGAREFDIFYOSEYELGRLDCTCRKR 728
661 VVCTFENENTEMCLAVWGMGAREFDIFYOSEYELGRLDCTCRKR 706

RESULT 5
US-09-833-245-1620
Sequence 1620, Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: 60/229, 358
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: 60/256, 931
 ; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/199, 384
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 2267
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO: 1620
 ; LENGTH: 706
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-833-245-1620

Query Match 97.1%; Score 3712; DB 22; Length 706;
 Best Local Similarity 100.0%; Pred. No. 1,1e-312;
 Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 23 MHAALQHPICIVALLIGISIHPLCFALDELAPLSLNTVLSENARDSSEFTPLGMLTOKIAYQ 82
DB 1 MHAALQHPICIVALLIGISIHPLCFALDELAPLSLNTVLSENARDSSEFTPLGMLTOKIAYQ 60

QY 83 IASGLAVLHKNNIIFCDLKSDNIIWMSLDVKEHINIKLSDYGISROSFHEGALGVESTPG 142
DB 61 IASGLAVLHKNNIIFCDLKSDNIIWMSLDVKEHINIKLSDYGISROSFHEGALGVESTPG 120

QY 143 YQAPRIPIRYDEKVDMEFYGMVLYELLSGSRPALGHQLOIAKKLSKIRPVLAGOPEE 202
DB 121 YQAPRIPIRYDEKVDMEFYGMVLYELLSGSRPALGHQLOIAKKLSKIRPVLAGOPEE 180

QY 203 VQFRRLQALMECWDTPKPRKPRPLALSVSQMKDPTFATFMYELCCGKQTAFFSSGOGEYT 262
DB 181 VQFRRLQALMECWDTPKPRKPRPLALSVSQMKDPTFATFMYELCCGKQTAFFSSGOGEYT 240

QY 263 VVFMDEKESRNTVYNTTEKGLMEVOVMCCPGMKVSCQLOVORSIMTATDOKIYYITLK 322
DB 241 VVFMDEKESRNTVYNTTEKGLMEVOVMCCPGMKVSCQLOVORSIMTATDOKIYYITLK 300

QY 323 GMPCLNTPOQALDTPAVVTCFLAVPIYIKKNSYLVLAGDLGVAVFPVYRGTPKDCSYL 382
DB 301 GMPCLNTPOQALDTPAVVTCFLAVPIYIKKNSYLVLAGDLGVAVFPVYRGTPKDCSYL 360

QY 383 CSHTANRSKFSIADDEARONPYVKAMEVYNSGSEWYSNGPGLVYIDCASLEICRLEP 442
DB 361 CSHTANRSKFSIADDEARONPYVKAMEVYNSGSEWYSNGPGLVYIDCASLEICRLEP 420

QY 443 YMAPSMVTSVYVCSSEGRGEVYVWCLDDKANSIWMYHSTYIQLCARFCGVPSPLRDMFPY 502
DB 421 YMAPSMVTSVYVCSSEGRGEVYVWCLDDKANSIWMYHSTYIQLCARFCGVPSPLRDMFPY 480

QY 503 RPLDTEPPAASHTANPKVPEGDSIADVSIWYSEELGTOLIIHOESLTDYCSMSYSSSP 562
DB 481 RPLDTEPPAASHTANPKVPEGDSIADVSIWYSEELGTOLIIHOESLTDYCSMSYSSSP 540

QY 563 RQARSPPSSLPSSPASSSVFSTDCEDSDMLTPGAASDRSEHDLTPMDGEFFSOHLQA 622
DB 541 RQARSPPSSLPSSPASSSVFSTDCEDSDMLTPGAASDRSEHDLTPMDGEFFSOHLQA 600

QY 623 VKLIAVRDLIWPFRGGDIYVIGLEKDESAORGRVIAVKARELTPHGVLYDAVAAYAKDT 682
DB 601 VKLIAVRDLIWPFRGGDIYVIGLEKDESAORGRVIAVKARELTPHGVLYDAVAAYAKDT 660

QY 683 VVCTFENENTEMCLAVWRMGAREFDIFYOSYEELGRLEACTRKR 728
DB 661 VVCTFENENTEMCLAVWRMGAREFDIFYOSYEELGRLEACTRKR 706
  
```

RESULT 6
 PCT-US01-01435-59
 ; Sequence 59, Application PC/TUS0101435
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc.
 ; TITLE OF INVENTION: 21 human secreted proteins

; FILE REFERENCE: PS726PCT
 ; CURRENT APPLICATION NUMBER: PCT/US01/01435
 ; CURRENT FILING DATE: 2001-01-17
 ; PRIOR APPLICATION NUMBER: 60/226, 282
 ; PRIOR FILING DATE: 2000-08-18
 ; NUMBER OF SEQ ID NOS: 145
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO: 59
 ; LENGTH: 706
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; PCT-US01-01435-59

Query Match 96.9%; Score 3705; DB 1; Length 706;
 Best Local Similarity 99.9%; Pred. No. 4,3e-312;
 Matches 705; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 23 MHAALQHPICIVALLIGISIHPLCFALDELAPLSLNTVLSENARDSSEFTPLGMLTOKIAYQ 82
DB 1 MHAALQHPICIVALLIGISIHPLCFALDELAPLSLNTVLSENARDSSEFTPLGMLTOKIAYQ 60

QY 83 IASGLAVLHKNNIIFCDLKSDNIIWMSLDVKEHINIKLSDYGISROSFHEGALGVESTPG 142
DB 61 IASGLAVLHKNNIIFCDLKSDNIIWMSLDVKEHINIKLSDYGISROSFHEGALGVESTPG 120

QY 143 YQAPRIPIRYDEKVDMEFYGMVLYELLSGSRPALGHQLOIAKKLSKIRPVLAGOPEE 202
DB 121 YQAPRIPIRYDEKVDMEFYGMVLYELLSGSRPALGHQLOIAKKLSKIRPVLAGOPEE 180

QY 203 VQFRRLQALMECWDTPKPRKPRPLALSVSQMKDPTFATFMYELCCGKQTAFFSSGOGEYT 262
DB 181 VQFRRLQALMECWDTPKPRKPRPLALSVSQMKDPTFATFMYELCCGKQTAFFSSGOGEYT 240

QY 263 VVFMDEKESRNTVYNTTEKGLMEVOVMCCPGMKVSCQLOVORSIMTATDOKIYYITLK 322
DB 241 VVFMDEKESRNTVYNTTEKGLMEVOVMCCPGMKVSCQLOVORSIMTATDOKIYYITLK 300

QY 323 GMPCLNTPOQALDTPAVVTCFLAVPIYIKKNSYLVLAGDLGVAVFPVYRGTPKDCSYL 382
DB 301 GMPCLNTPOQALDTPAVVTCFLAVPIYIKKNSYLVLAGDLGVAVFPVYRGTPKDCSYL 360

QY 383 CSHTANRSKFSIADDEARONPYVKAMEVYNSGSEWYSNGPGLVYIDCASLEICRLEP 442
DB 361 CSHTANRSKFSIADDEARONPYVKAMEVYNSGSEWYSNGPGLVYIDCASLEICRLEP 420

QY 443 YMAPSMVTSVYVCSSEGRGEVYVWCLDDKANSIWMYHSTYIQLCARFCGVPSPLRDMFPY 502
DB 421 YMAPSMVTSVYVCSSEGRGEVYVWCLDDKANSIWMYHSTYIQLCARFCGVPSPLRDMFPY 480

QY 503 RPLDTEPPAASHTANPKVPEGDSIADVSIWYSEELGTOLIIHOESLTDYCSMSYSSSP 562
DB 481 RPLDTEPPAASHTANPKVPEGDSIADVSIWYSEELGTOLIIHOESLTDYCSMSYSSSP 540

QY 563 RQARSPPSSLPSSPASSSVFSTDCEDSDMLTPGAASDRSEHDLTPMDGEFFSOHLQA 622
DB 541 RQARSPPSSLPSSPASSSVFSTDCEDSDMLTPGAASDRSEHDLTPMDGEFFSOHLQA 600

QY 623 VKLIAVRDLIWPFRGGDIYVIGLEKDESAORGRVIAVKARELTPHGVLYDAVAAYAKDT 682
DB 601 VKLIAVRDLIWPFRGGDIYVIGLEKDESAORGRVIAVKARELTPHGVLYDAVAAYAKDT 660

QY 683 VVCTFENENTEMCLAVWRMGAREFDIFYOSYEELGRLEACTRKR 728
DB 661 VVCTFENENTEMCLAVWRMGAREFDIFYOSYEELGRLEACTRKR 706
  
```

RESULT 7
 PCT-US01-11988-1621
 ; Sequence 1621, Application PC/TUS0111988
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc.
 ; TITLE OF INVENTION: Albumin Fusion Proteins
 ; FILE REFERENCE: PF546PCT

CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1621
LENGTH: 706
TYPE: PRT
ORGANISM: Homo sapiens
PCT: US01-11988-1621

Query Match 96.9%; Score 3705; DB 1; Length 706;
Best Local Similarity 99.9%; Pred. No. 4.3e-312;
Matches 705; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

23 MLHALQHPICVIALIGISHPCLFALELAPLSLNTVISENARDSSFTPLGHMLTOKIAYQ 82
1 MLHALQHPICVIALIGISHPCLFALELAPLSLNTVISENARDSSFTPLGHMLTOKIAYQ 60
83 IASGLAYLHKNNITFCDLKSDNIIWMSLDVKEHINIKISDYGISROSFHEGALVEGTPG 142
61 IASGLAYLHKNNITFCDLKSDNIIWMSLDVKEHINIKISDYGISROSFHEGALVEGTPG 120
143 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 202
121 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 180
203 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 262
181 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 240
263 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 322
241 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 300
323 GMCPLEPQOALDPAVVTCEFLAVPVIRKNSYLVLAGLADGLVAVFPVVRGTPKDCSYL 382
301 GMCPLEPQOALDPAVVTCEFLAVPVIRKNSYLVLAGLADGLVAVFPVVRGTPKDCSYL 360
363 CSHTANRKFSLADEDAQNIPYKAMEVNSGSEWYVSNBGLLVYIDCASLEICRLEP 442
361 CSHTANRKFSLADEDAQNIPYKAMEVNSGSEWYVSNBGLLVYIDCASLEICRLEP 420
443 YMAPSMVTSSVCSSEGRGEEVYVWCLDDKANSIWMYHSTTYOLCARYFCGVSPPLDMPV 502
421 YMAPSMVTSSVCSSEGRGEEVYVWCLDDKANSIWMYHSTTYOLCARYFCGVSPPLDMPV 480
503 RPLDEPAAHTANPKYPEGSDIADVSIMYSEELGTQILIHQESLTYCSMSYSSPP 562
481 RPLDEPAAHTANPKYPEGSDIADVSIMYSEELGTQILIHQESLTYCSMSYSSPP 540
563 ROAASPSLSPPASSSSVPESTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQHLQA 622
541 ROAASPSLSPPASSSSVPESTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQHLQA 600
623 VKIILAVRLIWPVRGGVIVYIGLEKDSGAQGRYIAVLAKEAELPHGVLDAAVYADT 682
601 VKIILAVRLIWPVRGGVIVYIGLEKDSGAQGRYIAVLAKEAELPHGVLDAAVYADT 660
683 VVCFEENETEMCLAVWRMGAREDFIYQSYEEELGRLEACTRKR 728
661 VVCFEENETEMCLAVWRMGAREDFIYQSYEEELGRLEACTRKR 706

RESULT 8
US-09-833-245-1621
Sequence 1621, Application US/09833245
GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: P546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1621
LENGTH: 706
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-1621

Query Match 96.9%; Score 3705; DB 22; Length 706;
Best Local Similarity 99.9%; Pred. No. 4.3e-312;
Matches 705; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

23 MLHALQHPICVIALIGISHPCLFALELAPLSLNTVISENARDSSFTPLGHMLTOKIAYQ 82
1 MLHALQHPICVIALIGISHPCLFALELAPLSLNTVISENARDSSFTPLGHMLTOKIAYQ 60
83 IASGLAYLHKNNITFCDLKSDNIIWMSLDVKEHINIKISDYGISROSFHEGALVEGTPG 142
61 IASGLAYLHKNNITFCDLKSDNIIWMSLDVKEHINIKISDYGISROSFHEGALVEGTPG 120
143 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 202
121 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 180
203 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 262
181 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 240
263 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 322
241 YQAEIRPRIYDEKVMFSTGMVLYELLSGORPALGHHOIAKKSIGRIPVIGQPEE 300
323 GMCPLEPQOALDPAVVTCEFLAVPVIRKNSYLVLAGLADGLVAVFPVVRGTPKDCSYL 382
301 GMCPLEPQOALDPAVVTCEFLAVPVIRKNSYLVLAGLADGLVAVFPVVRGTPKDCSYL 360
363 CSHTANRKFSLADEDAQNIPYKAMEVNSGSEWYVSNBGLLVYIDCASLEICRLEP 442
361 CSHTANRKFSLADEDAQNIPYKAMEVNSGSEWYVSNBGLLVYIDCASLEICRLEP 420
443 YMAPSMVTSSVCSSEGRGEEVYVWCLDDKANSIWMYHSTTYOLCARYFCGVSPPLDMPV 502
421 YMAPSMVTSSVCSSEGRGEEVYVWCLDDKANSIWMYHSTTYOLCARYFCGVSPPLDMPV 480
503 RPLDEPAAHTANPKYPEGSDIADVSIMYSEELGTQILIHQESLTYCSMSYSSPP 562
481 RPLDEPAAHTANPKYPEGSDIADVSIMYSEELGTQILIHQESLTYCSMSYSSPP 540
563 ROAASPSLSPPASSSSVPESTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQHLQA 622
541 ROAASPSLSPPASSSSVPESTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQHLQA 600
623 VKIILAVRLIWPVRGGVIVYIGLEKDSGAQGRYIAVLAKEAELPHGVLDAAVYADT 682
601 VKIILAVRLIWPVRGGVIVYIGLEKDSGAQGRYIAVLAKEAELPHGVLDAAVYADT 660
683 VVCFEENETEMCLAVWRMGAREDFIYQSYEEELGRLEACTRKR 728
661 VVCFEENETEMCLAVWRMGAREDFIYQSYEEELGRLEACTRKR 706

RESULT 9

```

US-10-132-382-6
; Sequence 6, Application US/10132382
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 6
; LENGTH: 1987
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-132-382-6

Query Match          93.5%; Score 3574.5; DB 25; Length 1987;
Best Local Similarity 94.5%; Pred. No. 4.9e-300;
Matches 687; Conservative 0; Mismatches 1; Indels 39; Gaps 1;

QY 1 MNRHLRATDAMKNFSEFROEASMLHALQHPCTIALIGISHPICFALTELAPLSLNTVLS 60
DB 1300 MNRHLRATDAMKNFSEFROEASMLHALQHPCTIALIGISHPICFALTELAPLSLNTVLS 1359
QY 61 ENARDSSEFIPGLHMLQKTAIYQASGLAYLHKNNIIFCDLKSNNILWMSLDVKEHINIKL 120
DB 1360 ENARDSSEFIPGLHMLQKTAIYQASGLAYLHKNNIIFCDLKSNNILWMSLDVKEHINIKL 1419
QY 121 SDYGISROSFHEGALVEGTPGYQAPDEIRPRIYDEKVMFSGYMWLYELLSGORPALGH 180
DB 1420 SDYGISROSFHEGALVEGTPGYQAPDEIRPRIYDEKVMFSGYMWLYELLSGORPALGH 1479
QY 181 HQQIAKLSKIGIRPLVGGPEEYQFRRLQALMECHDTPREKPRPLSVYSQMKDPTFAT 240
DB 1480 HQQIAKLSKIGIRPLVGGPEEYQFRRLQALMECHDTPREKPRPLSVYSQMKDPTFAT 1339
QY 241 FMYELCCGKOTAFSSQGGIYVFWDGKEESNNTYVNTTEKGLMEVORMCPCGMKVSQ 300
DB 1540 FMYELCCGKOTAFSSQGGIYVFWDGKEESNNTYVNTTEKGLMEVORMCPCGMKVSQ 1599
QY 301 LOVORSMTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLANPVIKKNSYLVLAGL 360
DB 1600 LOVORSMTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLANPVIKKNSYLVLAGL 1620
QY 361 ADGLVAVFVYVNGTPTDSCSYLCSHTANRKSFIADEDARONPYPYKAMEVYVNSGEVWY 420
DB 1621 ADGLVAVFVYVNGTPTDSCSYLCSHTANRKSFIADEDARONPYPYKAMEVYVNSGEVWY 1680
QY 421 SNGPGLLVIDCASLEICRLEPEYMAPSWNTSVYCSSEGGEEVWCLDKANSLVYHST 480
DB 1681 SNGPGLLVIDCASLEICRLEPEYMAPSWNTSVYCSSEGGEEVWCLDKANSLVYHST 1740
QY 481 TYQLCARFYCGVPSPLRDMFPVAPPLDTEPPAASHTANPVPEBDSIADVSIMYSEELGTQ 540
DB 1741 TYQLCARFYCGVPSPLRDMFPVAPPLDTEPPAASHTANPVPEBDSIADVSIMYSEELGTQ 1800
QY 541 IILHOSLTDYCSMSYSSSPPROARSPSLPSSPASSSSVPEFTDCEDSDMLHTPGAA 600
DB 1801 IILHOSLTDYCSMSYSSSPPROARSPSLPSSPASSSSVPEFTDCEDSDMLHTPGAA 1860
QY 601 SDRSEHDLTPMDGETFSQHLQAVKILAVRDLIMVPRGGDIVYIGLEKDSGAORGVIANV 660
DB 1861 SDRSEHDLTPMDGETFSQHLQAVKILAVRDLIMVPRGGDIVYIGLEKDSGAORGVIANV 1920
QY 661 LKARELTPHGVLYDAVAVKDTVYCFEENENTEMCLAVRGGARREFDIFYOSYEBLGRU 720
DB 1921 LKARELTPHGVLYDAVAVKDTVYCFEENENTEMCLAVRGGARREFDIFYOSYEBLGRU 1980
QY 721 EACTRRK 727
DB 1981 EACTRRK 1987

```

```

RESULT 10
US-10-132-382-2
; Sequence 2, Application US/10132382
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 2013
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-132-382-2

Query Match          93.5%; Score 3574.5; DB 25; Length 2013;
Best Local Similarity 94.5%; Pred. No. 5e-300;
Matches 687; Conservative 0; Mismatches 1; Indels 39; Gaps 1;

QY 1 MNRHLRATDAMKNFSEFROEASMLHALQHPCTIALIGISHPICFALTELAPLSLNTVLS 60
DB 1326 MNRHLRATDAMKNFSEFROEASMLHALQHPCTIALIGISHPICFALTELAPLSLNTVLS 1385
QY 61 ENARDSSEFIPGLHMLQKTAIYQASGLAYLHKNNIIFCDLKSNNILWMSLDVKEHINIKL 120
DB 1386 ENARDSSEFIPGLHMLQKTAIYQASGLAYLHKNNIIFCDLKSNNILWMSLDVKEHINIKL 1445
QY 121 SDYGISROSFHEGALVEGTPGYQAPDEIRPRIYDEKVMFSGYMWLYELLSGORPALGH 180
DB 1446 SDYGISROSFHEGALVEGTPGYQAPDEIRPRIYDEKVMFSGYMWLYELLSGORPALGH 1505
QY 181 HQQIAKLSKIGIRPLVGGPEEYQFRRLQALMECHDTPREKPRPLSVYSQMKDPTFAT 240
DB 1506 HQQIAKLSKIGIRPLVGGPEEYQFRRLQALMECHDTPREKPRPLSVYSQMKDPTFAT 1565
QY 241 FMYELCCGKOTAFSSQGGIYVFWDGKEESNNTYVNTTEKGLMEVORMCPCGMKVSQ 300
DB 1566 FMYELCCGKOTAFSSQGGIYVFWDGKEESNNTYVNTTEKGLMEVORMCPCGMKVSQ 1625
QY 301 LOVORSMTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLANPVIKKNSYLVLAGL 360
DB 1626 LOVORSMTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLANPVIKKNSYLVLAGL 1646
QY 361 ADGLVAVFVYVNGTPTDSCSYLCSHTANRKSFIADEDARONPYPYKAMEVYVNSGEVWY 420
DB 1647 ADGLVAVFVYVNGTPTDSCSYLCSHTANRKSFIADEDARONPYPYKAMEVYVNSGEVWY 1706
QY 421 SNGPGLLVIDCASLEICRLEPEYMAPSWNTSVYCSSEGGEEVWCLDKANSLVYHST 480
DB 1707 SNGPGLLVIDCASLEICRLEPEYMAPSWNTSVYCSSEGGEEVWCLDKANSLVYHST 1766
QY 481 TYQLCARFYCGVPSPLRDMFPVAPPLDTEPPAASHTANPVPEBDSIADVSIMYSEELGTQ 540
DB 1767 TYQLCARFYCGVPSPLRDMFPVAPPLDTEPPAASHTANPVPEBDSIADVSIMYSEELGTQ 1826
QY 541 IILHOSLTDYCSMSYSSSPPROARSPSLPSSPASSSSVPEFTDCEDSDMLHTPGAA 600
DB 1827 IILHOSLTDYCSMSYSSSPPROARSPSLPSSPASSSSVPEFTDCEDSDMLHTPGAA 1886
QY 601 SDRSEHDLTPMDGETFSQHLQAVKILAVRDLIMVPRGGDIVYIGLEKDSGAORGVIANV 660
DB 1887 SDRSEHDLTPMDGETFSQHLQAVKILAVRDLIMVPRGGDIVYIGLEKDSGAORGVIANV 1946
QY 661 LKARELTPHGVLYDAVAVKDTVYCFEENENTEMCLAVRGGARREFDIFYOSYEBLGRU 720
DB 1947 LKARELTPHGVLYDAVAVKDTVYCFEENENTEMCLAVRGGARREFDIFYOSYEBLGRU 2006
QY 721 EACTRRK 727
DB 2007 EACTRRK 2013

```

RESULT 11
 US-10-132-382-8
 ; Sequence 8, Application US/10132382
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 8
 ; LENGTH: 2014
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-132-382-8

Query Match 93.5%; Score 3574.5; DB 25; Length 2014;
 Best Local Similarity 94.5%; Pred. No. 5e-300;
 Matches 687; Conservative 0; Mismatches 1; Indels 39; Gaps 1;

QY 1 MLRHRTADAKNFSEFROEASMLALHPCTVALIGISHPCLFALAPLSSINTVLS 60
 DB 1337 MLRHRTADAKNFSEFROEASMLALHPCTVALIGISHPCLFALAPLSSINTVLS 1386
 QY 61 ENARDSFPIPLGHMLTKIAYOIASGLAYLHKNNIFCDLSKDNILVMSLDVKEHINTKL 120
 DB 1387 ENARDSFPIPLGHMLTKIAYOIASGLAYLHKNNIFCDLSKDNILVMSLDVKEHINTKL 1446
 QY 121 SDYGISROSFHGALGVGTGTCYQAPETRPRIYDEKDMSEYGVLYELLSGORPALGH 180
 DB 1447 SDYGISROSFHGALGVGTGTCYQAPETRPRIYDEKDMSEYGVLYELLSGORPALGH 1506
 QY 181 HOLOIAKLSKIRPYLQOPEEVOFRLOALMECHDKRPERPLALSVSOMKPTFAT 240
 DB 1507 HOLOIAKLSKIRPYLQOPEEVOFRLOALMECHDKRPERPLALSVSOMKPTFAT 1566
 QY 241 FMYELCCGKQTAFFSSQOGEYTVFMDGKEESRNTYVNTKGLMEVOBMCCPGMKVSCQ 300
 DB 1567 FMYELCCGKQTAFFSSQOGEYTVFMDGKEESRNTYVNTKGLMEVOBMCCPGMKVSCQ 1626
 QY 301 LOVOSLTATDOKIYITTLKGMCPMLNTPQALDTPAVVTCFLAVPYIKKSYVYIAGL 360
 DB 1627 LOVOSLTATDOKIYITTLKGMCPMLNTPQALDTPAVVTCFLAVPYIKKSYVYIAGL 1647
 QY 361 ADGLVAFVYVNGTFRKDCSYLCSHTANRKSFIADEDARONPYKAMEVNSGSEVY 420
 DB 1648 ADGLVAFVYVNGTFRKDCSYLCSHTANRKSFIADEDARONPYKAMEVNSGSEVY 1707
 QY 421 SNGPGLVIDCASLEICRLEBYMAPSWMTSVYCSSEGESEVWVCLDOKANSLVYHST 480
 DB 1708 SNGPGLVIDCASLEICRLEBYMAPSWMTSVYCSSEGESEVWVCLDOKANSLVYHST 1767
 QY 481 TYOLCARFYCGVPSPLRDMFVPRPLDTEPPAASHANPVPBGDSIADVSIMYSEELGQ 540
 DB 1768 TYOLCARFYCGVPSPLRDMFVPRPLDTEPPAASHANPVPBGDSIADVSIMYSEELGQ 1827
 QY 541 ILIHOSLTIDYCSMSYSSPPROAARSPSSLPSPASSSVFSTDCDSMLHTPGA 600
 DB 1828 ILIHOSLTIDYCSMSYSSPPROAARSPSSLPSPASSSVFSTDCDSMLHTPGA 1887
 QY 601 SDRSEHDLTPMDGETFSQHLQA VKILAVRDLIMVPRRGDVIVYIGLEKDSGQORGVIAV 660
 DB 1888 SDRSEHDLTPMDGETFSQHLQA VKILAVRDLIMVPRRGDVIVYIGLEKDSGQORGVIAV 1947
 QY 661 LKARELTPIHGVLYDAVAVKDVTCTFENENTEMCLAVRGAGARFDFIYOSYBELGL 720
 DB 1948 LKARELTPIHGVLYDAVAVKDVTCTFENENTEMCLAVRGAGARFDFIYOSYBELGL 2007
 QY 721 EACTRR 727
 DB 2008 EACTRR 2014

RESULT 12
 US-10-132-382-4
 ; Sequence 4, Application US/10132382
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 4
 ; LENGTH: 2040
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-132-382-4

Query Match 93.5%; Score 3574.5; DB 25; Length 2040;
 Best Local Similarity 94.5%; Pred. No. 5.1e-300;
 Matches 687; Conservative 0; Mismatches 1; Indels 39; Gaps 1;

QY 1 MLRHRTADAKNFSEFROEASMLALHPCTVALIGISHPCLFALAPLSSINTVLS 60
 DB 1333 MLRHRTADAKNFSEFROEASMLALHPCTVALIGISHPCLFALAPLSSINTVLS 1412
 QY 61 ENARDSFPIPLGHMLTKIAYOIASGLAYLHKNNIFCDLSKDNILVMSLDVKEHINTKL 120
 DB 1413 ENARDSFPIPLGHMLTKIAYOIASGLAYLHKNNIFCDLSKDNILVMSLDVKEHINTKL 1472
 QY 121 SDYGISROSFHGALGVGTGTCYQAPETRPRIYDEKDMSEYGVLYELLSGORPALGH 180
 DB 1473 SDYGISROSFHGALGVGTGTCYQAPETRPRIYDEKDMSEYGVLYELLSGORPALGH 1532
 QY 181 HOLOIAKLSKIRPYLQOPEEVOFRLOALMECHDKRPERPLALSVSOMKPTFAT 240
 DB 1533 HOLOIAKLSKIRPYLQOPEEVOFRLOALMECHDKRPERPLALSVSOMKPTFAT 1592
 QY 241 FMYELCCGKQTAFFSSQOGEYTVFMDGKEESRNTYVNTKGLMEVOBMCCPGMKVSCQ 300
 DB 1593 FMYELCCGKQTAFFSSQOGEYTVFMDGKEESRNTYVNTKGLMEVOBMCCPGMKVSCQ 1652
 QY 301 LOVOSLTATDOKIYITTLKGMCPMLNTPQALDTPAVVTCFLAVPYIKKSYVYIAGL 360
 DB 1653 LOVOSLTATDOKIYITTLKGMCPMLNTPQALDTPAVVTCFLAVPYIKKSYVYIAGL 1673
 QY 361 ADGLVAFVYVNGTFRKDCSYLCSHTANRKSFIADEDARONPYKAMEVNSGSEVY 420
 DB 1674 ADGLVAFVYVNGTFRKDCSYLCSHTANRKSFIADEDARONPYKAMEVNSGSEVY 1733
 QY 421 SNGPGLVIDCASLEICRLEBYMAPSWMTSVYCSSEGESEVWVCLDOKANSLVYHST 480
 DB 1734 SNGPGLVIDCASLEICRLEBYMAPSWMTSVYCSSEGESEVWVCLDOKANSLVYHST 1793
 QY 481 TYOLCARFYCGVPSPLRDMFVPRPLDTEPPAASHANPVPBGDSIADVSIMYSEELGQ 540
 DB 1794 TYOLCARFYCGVPSPLRDMFVPRPLDTEPPAASHANPVPBGDSIADVSIMYSEELGQ 1853
 QY 541 ILIHOSLTIDYCSMSYSSPPROAARSPSSLPSPASSSVFSTDCDSMLHTPGA 600
 DB 1854 ILIHOSLTIDYCSMSYSSPPROAARSPSSLPSPASSSVFSTDCDSMLHTPGA 1913
 QY 601 SDRSEHDLTPMDGETFSQHLQA VKILAVRDLIMVPRRGDVIVYIGLEKDSGQORGVIAV 660
 DB 1914 SDRSEHDLTPMDGETFSQHLQA VKILAVRDLIMVPRRGDVIVYIGLEKDSGQORGVIAV 1973
 QY 661 LKARELTPIHGVLYDAVAVKDVTCTFENENTEMCLAVRGAGARFDFIYOSYBELGL 720
 DB 1974 LKARELTPIHGVLYDAVAVKDVTCTFENENTEMCLAVRGAGARFDFIYOSYBELGL 2033
 QY 721 EACTRR 727
 DB 2033 EACTRR 2033

Db 2034 EACTRR 2040

RESULT 13

PCT-US01-08631-38410
Sequence 38410, Application PC/TUS0108631

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc

TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 21272-049

CURRENT APPLICATION NUMBER: PCT/US01/08631

PRIOR FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: 09/540,217

PRIOR FILING DATE: 2000-03-31

PRIOR APPLICATION NUMBER: 09/649,167

PRIOR FILING DATE: 2000-08-23

NUMBER OF SEQ ID NOS: 60736

SOFTWARE: Custom

SEQ ID NO 38410

LENGTH: 809

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: DOMAIN

LOCATION: (223)..(261)

OTHER INFORMATION: Receptor tyrosine kinase class III proteins domain identified

OTHER INFORMATION: by EMATRIX, accession number BL00240E, p-value=7.660e-10, raw

OTHER INFORMATION: score of 11.56

NAME/KEY: DOMAIN

LOCATION: (103)..(372)

OTHER INFORMATION: Eukaryotic protein kinase domain identified by Pfam,

OTHER INFORMATION: accession name pkinase, E-value=1.8e-44, Pfam score of 161.2

NAME/KEY: misc_feature

LOCATION: (1)..(809)

OTHER INFORMATION: xaa - X or * as defined in Table 2

PCT-US01-08631-38410

Query Match

Best Local Similarity 78.4%; Score 2997; DB 1; Length 809;

Matches 58; Conservative 4; Mismatches 26; Indels 12; Gaps 4;

Db 1 MLRLRLATDAMKNSFSEFROASHMLALHPCIVALLIGISIHPLCEALPLSLNLYLS 60
Db 149 MLRLRLATDAMKNSFSEFROASHMLALHPCIVALLIGISIHPLCEALPLSLNLYLS 208
QY 61 ENARSSFIPLGMLTOKIAYQIASGLAYLHKNNIIFCDKSDNIIYMSLDVKEHINIKL 120
Db 209 ENARSSFIPLGMLTOKIAYQIASGLAYLHKNNIIFCDKSDNIIYMSLDVKEHINIKL 268
QY 121 SDYGISNOSHFGALVGEFGPGYQAPFIRIYDEKVDMSFGVLYELLSCGPALGH 180
Db 269 SDYGISNOSHFGALVGEFGPGYQAPFIRIYDEKVDMSFGVLYELLSCGPALGH 328
QY 181 HQQIAKKLSKIGIRPVLGQPEEVOFRRLQALMECHMDTKREKRLALSVYSOMKAPDFE-- 238
Db 329 HQQIAKKLSKIGIRPVLGQPEEVOFRRLQALMECHMDTKREKRLALSVYSOMKAPDFE 388
QY 239 ---ATNRYELCGCKQTAFFSSQGOEYTVFWD--GKEESRNYTVNTEKGLMEYORMCCPG 294
Db 389 LHYXTVLMW---ADSLILIPGVRHSGVLRGKMGVCELHG---GNTREKGLMEYORMCCPG 442
QY 295 MKVSCQLOVORSLMTATEODKIYITLLKGMCLPNTPOQALDTPAVYTCFLAVYIKNSY 354
Db 443 MKVSCQLOVORSLMTATEODKIYITLLKGMCLPNTPOQALDTPAVYTCFLAVYIKNSY 502
QY 355 LVYAGLADGLVAVFPVVRGTPKDCSYLCSHTANRSKFSIADBDARONPYPVAMEVYNS 414
Db 503 LVYAGLADGLVAVFPVVRGTPKDCSYLCSHTANRSKFSIADBDARONPYPVAMEVYNS 562
QY 415 GSEVWTSNPGGLVITOCASLEICRLEPYPAPSMVTSVYVCSSEGRGEEVYVCLDDKANS 474
Db 563 GSEVWTSNPGGLVITOCASLEICRLEPYPAPSMVTSVYVCSSEGRGEEVYVCLDDKANS 622

QY 475 VMHSTTYOLCARFCGVPSPFLRDMFPVRLDTEPPAASHANKVPEGSDIADVSIMYS 534
Db 623 VMHSTTYOLCARFCGVPSPFLRDMFPVRLDTEPPAASHANKVPEGSDIADVSIMYS 682
QY 535 EELGTOLLIHQESTLDYCSMSYSSSPPROAARSPSSLPSPASSSSVPSTDCEDSDML 594
Db 683 EELGTOLLIHQESTLDYCSMSYSSSPPROAARSPSSLPSPASSSSVPSTDCEDSDML 742
QY 595 HTPGAASDRSEHDLTPMDGETFSOHL 620
Db 743 HTPGAASDRSEHDLTPMDGETFSOHL 768

RESULT 14

PCT-US02-05109-982

Sequence 982, Application PC/TUS0205109

GENERAL INFORMATION:

APPLICANT: Hyseq Inc.

APPLICANT: Tang, Tom Y.

APPLICANT: Zhou, Ping

APPLICANT: Goodrich, Ryle

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Zhao, Qing A.

APPLICANT: Ren, Feiyan

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Ma, Yunding

APPLICANT: Yamazaki, Victoria

APPLICANT: Chen, Rui-hong

APPLICANT: Wang, Zhwei

APPLICANT: Ghosh, Malabika

APPLICANT: Wehrman, Tom

APPLICANT: Wang, Jian-Rui

APPLICANT: Wang, Dunrui

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 21272-094-061

CURRENT APPLICATION NUMBER: PCT/US02/05109

PRIOR FILING DATE: 2002-03-14

PRIOR APPLICATION NUMBER: 09/810,173

PRIOR FILING DATE: 2001-03-15

NUMBER OF SEQ ID NOS: 1052

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 982

LENGTH: 320

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US02-05109-982

Query Match

Best Local Similarity 43.9%; Score 1679; DB 1; Length 320;

Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 409 MEVYNSGSEVWYSNGPGLVIDCASLEICRLEPYPAPSMVTSVYVCSSEGRGEEVYVCLD 468
Db 1 MEVYNSGSEVWYSNGPGLVIDCASLEICRLEPYPAPSMVTSVYVCSSEGRGEEVYVCLD 60
QY 469 DKANSIYVHSTTYOLCARFCGVPSPFLRDMFPVRLDTEPPAASHANKVPEGSDIAD 538
Db 61 DKANSIYVHSTTYOLCARFCGVPSPFLRDMFPVRLDTEPPAASHANKVPEGSDIAD 120
QY 529 VSIYSEBELGTOLLIHQESTLDYCSMSYSSSPPROAARSPSSLPSPASSSSVPSTDC 588
Db 121 VSIYSEBELGTOLLIHQESTLDYCSMSYSSSPPROAARSPSSLPSPASSSSVPSTDC 180
QY 589 ESDDMHTPGAASDRSEHDLTPMDGETFSOHLQAVKILAVRDLIWPVRGGDVYVIGLEK 648
Db 181 ESDDMHTPGAASDRSEHDLTPMDGETFSOHLQAVKILAVRDLIWPVRGGDVYVIGLEK 740
QY 649 DSEAORGRVYAVKARELPYHGVLYDAAYVAKDTVCTEENENTENCILAVWVGWGRRED 708
Db 241 DSEAORGRVYAVKARELPYHGVLYDAAYVAKDTVCTEENENTENCILAVWVGWGRRED 800

QY 709 IFYOSYEELGRLACTRRR 728
DB 301 IFYOSYEELGRLACTRRR 320

RESULT 15

US-60-155-811-774
; Sequence 774, Application US/60155811
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vito
; TITLE OF INVENTION: Isolated GPCR proteins, nucleic acid
; FILE REFERENCE: CLO00095
; CURRENT APPLICATION NUMBER: US/60/155,811
; CURRENT FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 1060
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 774
; LENGTH: 1050
; TYPE: PRT
; ORGANISM: Drosophila
US-60-155-811-774

Query Match 11.4%; Score 435; DB 27; Length 1050;
Best Local Similarity 25.3%; Pred. No. 1.5e-27;
Matches 170; Conservative 103; Mismatches 254; Indels 146; Gaps 26;

QY 2 LRLRLATDAMKNSFROEASMLHLOHPCIVALLIGISTHPLCFALPLSLNTVLSL 61
DB 397 LQH-----SKATCTARQELAVLTILKHPNIVPLVGICIKPLALVELAPLGGLDLRLH 451
QY 62 NARDSFIPGLHMLQKIAVQIASGLAYLHKNIIFCDLKSNDILVMSL-----DYKE 114
DB 452 YRRSGAH--MGPHPTQTLVQAARALEYLRRIIYRDLKSENVLYWELPQPTEDSPRN 509
QY 115 HINIKISDYGISROSHFEGALVEGTPGYQAPET---RPRIYDEKVMESYGMVLYELL 171
DB 510 LVHIKIDADGISKQFAPGAKGFGTEGFMPEIIRYNGEETERYKDFCSFGMFIYENI 569
QY 172 SCGRPALGHQLOIAKLKSGIRPVLGQPEVQFR--LQALMECWDTKPERRLASVY 230
DB 570 SLRQPEGHE--SIKECILGSSRLPTQ--RETQPTCCLDLWLCWHEQPRRRPTASQIV 626
QY 231 SQMKDPTFAFM-----YELCGCK--QTAFSSQGEYTVFMDEKESRNYTVNTE 281
DB 627 SLSABECIHLDVAMPHEKIVCGVQSLVGMGDERCGLEMLPFSGRIDIIDCSP 686
QY 282 KG-LMEVQRMCC-----PQMKVSCQLOVQRLMTATEQ 314
DB 687 SGLLQCNISCSPPQVAPKTPENGANSRANSQRLPRMNLCCCLVGEALIMGDVSG 746
QY 315 KIITYTLKGCPLNTPOQALDTPAVTCTFLAVVIKNSLVLAGLADGLVAFVYVRCGT 374
DB 747 NLHAYSTSTYAHLS--YMLD--PNIKSAVISLVYMERIA--RAVVGTHNG--RVELVDATQ 800
QY 375 PKDSCSY-----ICSHTANRSKSTIADED-----ARQNPYP----- 405
DB 801 MBSNCAFAEGSFVLTETCSGFVLAACSVYVDGTYELMCGETAGKINVPFLNENGVSCHQ 860
QY 406 -----VKAMEVYVNGSEVYVNSGGLLV-----IDCASELTC--R 438
DB 861 ALCHSEPNLIDVYKVARMCNSHVFSCLYPCMYQWQVISKRIENKIDCSKILPCSE 920
QY 439 RLEPVYAPBSVYTVSSSGEGRBEVVMCLDDKANSLSVMYHSTIYQLCARFCGVPSPLRD 498
DB 921 SLQSLAIDEHVNLIK-----QISALAHNSELY--IGTTWGLIVALEHT 964
QY 499 MFPV---RPLDTPPAASHANKVPEGDSIAVSIYSELTQILIHQESLTDYCSMS 555
DB 965 LRIYVFRPEYENIKSITISKDNP---LIATIGRRY--RSLISRYVDSASST---KS 1016
QY 556 SYSSSPPROAARS 568

DB 1017 SAVSTPTHGAARS 1029

Search completed: April 8, 2003, 10:32:30
Job time : 333 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:23:49 ; Search time 78 Seconds
(without alignments)
1367.278 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822

Sequence: 1 MLRLRLATDANKNFSEFRQ.....IFYSYELGRLNACTRRKR 728

Scoring table:

BLOSUM62

Searched: 666312 seqs, 146494096 residues

Total number of hits satisfying chosen parameters: 666312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending Patents AA New *

- 1: /cgn2_6/ptodata/2/paa/PCF_NEM_COMB.pcp:**
- 2: /cgn2_6/ptodata/2/paa/US06_NEM_COMB.pcp:**
- 3: /cgn2_6/ptodata/2/paa/US07_NEM_COMB.pcp:**
- 4: /cgn2_6/ptodata/2/paa/US08_NEM_COMB.pcp:**
- 5: /cgn2_6/ptodata/2/paa/US09_NEM_COMB.pcp:**
- 6: /cgn2_6/ptodata/2/paa/US10_NEM_COMB.pcp:**
- 7: /cgn2_6/ptodata/2/paa/US60_NEM_COMB.pcp:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	3331	87.2	659	1	PCF-US03-01363-170
2	3253	86.0	501	6	US-10-094-749-1689
3	327.5	8.6	656	6	US-10-311-034-5
4	289.5	7.6	953	1	PCF-US03-04816-56
5	289.5	7.6	953	6	US-10-369-022-56
6	289.5	7.5	953	6	US-60-426-964-2
7	285	7.5	954	6	US-10-263-929-187
8	283.5	7.4	1002	6	US-10-263-929-187
9	273.5	7.2	1081	1	PCF-US02-39126-13
10	273.5	7.2	1096	6	US-10-210-130-110
11	271	7.1	251	6	US-10-052-648A-42
12	271	7.1	254	6	US-10-074-978A-151
13	271	7.1	256	6	US-10-072-012-800
14	271	7.1	256	6	US-10-072-012-857
15	271	7.1	256	6	US-10-087-684-99
16	270	7.1	1046	6	US-10-258-106-9
17	268.5	7.0	438	6	US-10-361-942-51
18	268	7.0	249	6	US-10-074-978A-162
19	267.5	7.0	1097	6	US-10-288-798-12
20	267.5	7.0	1097	6	US-10-362-892-12
21	265	6.9	732	6	US-10-310-154-354
22	263	6.9	642	6	US-60-443-566-4543
23	263	6.9	964	6	US-10-210-130-32
24	263	6.9	1036	1	PCF-US02-34037-2
25	263	6.9	1036	1	PCF-US03-02568-24
26	263	6.9	1036	6	US-10-263-929-189

27	263	6.9	1036	6	US-10-354-358-24	Sequence 24, Appl
28	260.5	6.8	847	5	US-09-949-016-6222	Sequence 6222, Ap
29	260.5	6.8	847	6	US-10-263-929-188	Sequence 188, App
30	260.5	6.8	847	6	US-10-348-119-327	Sequence 327, App
31	260.5	6.8	885	5	US-09-949-016-7789	Sequence 7789, Ap
32	260	6.8	511	5	US-09-724-676-68428	Sequence 68428, A
33	260	6.8	511	5	US-09-724-676-68428	Sequence 68428, A
34	260	6.8	610	7	US-60-443-566-3046	Sequence 3046, Ap
35	260	6.8	610	7	US-60-452-680-19069	Sequence 19069, A
36	260	6.8	663	3	US-09-949-016-7198	Sequence 7198, Ap
37	259.5	6.8	850	6	US-10-263-929-192	Sequence 192, App
38	247.5	6.5	922	6	US-10-263-929-190	Sequence 190, App
39	246.5	6.4	765	6	US-10-285-325-2	Sequence 2, Appl
40	244.5	6.4	823	6	US-10-376-564-63	Sequence 63, Appl
41	244	6.4	938	5	US-09-724-676-68429	Sequence 68429, A
42	244	6.4	938	5	US-09-724-676A-68429	Sequence 68429, A
43	243.5	6.4	505	6	US-10-099-056-1244	Sequence 1244, Ap
44	243	6.4	455	6	US-10-094-749-2477	Sequence 2477, Ap
45	243	6.4	800	7	US-60-438-000-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
PCF-US03-01363-170
Sequence 170, Application PC/TUS0301363
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.; JONES, Anissa L.;
APPLICANT: DAHL, Christopher R.; GIETEN, Darryl;
APPLICANT: CHINN, Joyce; DUPOUR, Gerard E.;
APPLICANT: JACKSON, Jennifer L.; YU, Jimmy Y.;
APPLICANT: TUASON, Olivia; YAP, Pierre E.;
APPLICANT: AMSHEV, Stefan R.; DAM, Tam C.;
APPLICANT: LIU, Tommy F.; GERSTIN Jr., Edward H.;
APPLICANT: PERALTA, Carey H.; LEWIS, Samantha A.;
APPLICANT: CHEN, Alice; MARAHN, Rakeb;
APPLICANT: LAN, Ruth Y.; URASHKA, Michael;
APPLICANT: R., KRISTIAN, Sreenivasa R.; KOLLURU, Vijaykumar;
APPLICANT: PANESAR, Iqbal S.
TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
FILE REFERENCE: PT-1240 PCT
CURRENT APPLICATION NUMBER: PCF/US03/01363
PRIOR FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: US 60/349,946
PRIOR FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: US 60/349,413
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PERL Program
SEQ ID NO 170
LENGTH: 659
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID NO: LG:337452.25.orf3:2002JAN18
PCF-US03-01363-170

Query Match 87.2%; Score 3331; DB 1; Length 659;
Best Local Similarity 94.8%; Pred. No. 8e-233;
Matches 641; Conservative 6; Mismatches 9; Indels 20; Gaps 3;

QY 38 ISIHPLCFALTEAPLSLTVLSEANRDSFPLGLMLTOKIAYQIASGLAVLHKNIIF 97
DB 1 ISIHPLCFALTEAPLSLTVLSEANRDSFPLGLMLTOKIAYQIASGLAVLHKNIIF 60
QY 98 CDKSDNLIWSLVDYKEHINIKLSDYGISRQSFHEGALGVECTPGYQAEIRPRIYDER 157
DB 61 CDKSDNLIWSLVDYKEHINIKLSDYGISRQSFHEGALGVECTPGYQAEIRPRIYDER 120
QY 158 VMEFSYGMVLYLLSGQRALHHQIQIAKTKSGIRPLVPGQPEYQFRDLALMECHD 217

```
Db 121 VDMFSGVLYELLGSRPALGHQLOIAKKLSKIRPVLGQPEVOFRRLQALMECWD 180
QY 218 TKPEKRLALSVYQOMKDPFATFMYELCCGKQTAFFSSQGEYTVFMNGKEESRNTYV 277
Db 181 TKPEKRLALSVYQOMKDPFATFMYELCCGKQTAFFSSQGEYTVFMNGKEESRNTYV 240
QY 278 VTEKGLMEVQRMCCPGMKVSCOLQVQSRSLMTATEDOKIYITLKGMCPLNTPOQALDTP 337
Db 241 VTEKGLMEVQRMCCPGMKVSCOLQVQSRSLMTATEDOKIYITLKGMCPLNTPOQALDTP 300
QY 338 AVYTCFLAVPVYIKKNSYLVLAGLADGLVAVFPVYRGTPKDCSCYLSHTANRKSFTADE 397
Db 301 AVYTCFLAVPVYIKKNSYLVLAGLADGLVAVFPVYRGTPKDCSCYLSHTANRKSFTADE 360
QY 398 DARNOPPVKAMEVYVNSGSEWYVSNGLVITDCALEICRLEPYMAPSWTVSVYCSSE 457
Db 361 DARNOPPVKAMEVYVNSGSEWYVSNGLVITDCALEICRLEPYMAPSWTVSVYCSSE 420
QY 458 GRGEVYVWCLDDKANSLVYHSTTYQLCARVFCGVPSPRLDMFPVRLDTEPPAASHATAN 517
Db 421 GRGEVYVWCLDDKANSLVYHSTTYQLCARVFCGVPSPRLDMFPVRLDTEPPAASHATAN 480
QY 518 PKYPEGDSIADVSIYSEELGTQILIHQESLTDYCSMSYSSSPRQAARSPSSLPSSPA 577
Db 481 PKYPEGDSIADVSIYSEELGTQILIHQESLTDYCSMSYSSSPRQAARSPSSLPSSPA 540
QY 578 SSSVFPSTDCEDSDMLHTPGAASDRSEHDLTPMDGETFQSOHLQAVYILAVRDLIMVPRR 637
Db 541 SSSVFPSTDCEDSDMLHTPGAASDRSEHDLTPMDGETFQSOHLQAVYILAVRDLIMVPRR 600
QY 638 GGDVYVIGLEKDESEAGRYAVLAKARELTPHGLVDAVAVAKDVCTFENEMENTWCLA 697
Db 601 GGDVYVIGLEKDESEAGRYAVLAKARELTPHGLVDAVAVAKDVCTFENEMENTWCLA 643
QY 698 VMKGWAKREDIFQOS 713
Db 644 -WAGWMPVR--DMYMA 656

RESULT 2
US-10-094-749-1689
; Sequence 1689, Application US/10094749
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YUKI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: TOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1689
; LENGTH: 501
```

```
TYPE: PR1
ORGANISM: Homo sapiens
US-10-094-749-1689

Query Match 66.0%; Score 2523; DB 6; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.5e-189;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 MESYGMVLYELLGSRPALGHQLOIAKKLSKIRPVLGQPEVOFRRLQALMECWD 219
Db 1 MESYGMVLYELLGSRPALGHQLOIAKKLSKIRPVLGQPEVOFRRLQALMECWD 60
QY 220 PEKRLALSVYQOMKDPFATFMYELCCGKQTAFFSSQGEYTVFMNGKEESRNTYV 279
Db 61 PEKRLALSVYQOMKDPFATFMYELCCGKQTAFFSSQGEYTVFMNGKEESRNTYV 120
QY 280 TEKGLMEVQRMCCPGMKVSCOLQVQSRSLMTATEDOKIYITLKGMCPLNTPOQALDTPAV 339
Db 121 TEKGLMEVQRMCCPGMKVSCOLQVQSRSLMTATEDOKIYITLKGMCPLNTPOQALDTPAV 180
QY 340 VTGFLAVPVYIKKNSYLVLAGLADGLVAVFPVYRGTPKDCSCYLSHTANRKSFTADE 399
Db 181 VTGFLAVPVYIKKNSYLVLAGLADGLVAVFPVYRGTPKDCSCYLSHTANRKSFTADE 240
QY 400 ROPPYPKAMEVYVNSGSEWYVSNGLVITDCALEICRLEPYMAPSWTVSVYCSSEGR 459
Db 241 ROPPYPKAMEVYVNSGSEWYVSNGLVITDCALEICRLEPYMAPSWTVSVYCSSEGR 300
QY 460 GEEVYVWCLDDKANSLVYHSTTYQLCARVFCGVPSPRLDMFPVRLDTEPPAASHATAN 519
Db 301 GEEVYVWCLDDKANSLVYHSTTYQLCARVFCGVPSPRLDMFPVRLDTEPPAASHATAN 360
QY 520 VPEGDSIADVSIYSEELGTQILIHQESLTDYCSMSYSSSPRQAARSPSSLPSSPASS 579
Db 361 VPEGDSIADVSIYSEELGTQILIHQESLTDYCSMSYSSSPRQAARSPSSLPSSPASS 420
QY 580 SSVFPSTDCEDSDMLHTPGAASDRSEHDLTPMDGETFQSOHLQAVYILAVRDLIMVPRR 637
Db 421 SSVFPSTDCEDSDMLHTPGAASDRSEHDLTPMDGETFQSOHLQAVYILAVRDLIMVPRR 478

RESULT 3
US-10-311-034-5
; Sequence 5, Application US/10311034
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YOE, Henry
; APPLICANT: IAL, Preeti
; APPLICANT: BANDMAN, Olga
; APPLICANT: BOROMSKY, Mark L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBODLEY, Catherine M.
; APPLICANT: CHAMLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: GREENWALD, Sara R.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: KEARNEY, Liam
; APPLICANT: BURFORD, Neil
; APPLICANT: NGUYEN, Damiel B.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HE, Ann
; APPLICANT: THORNTON, Michael
; APPLICANT: HAFALIA, April
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: LO, Terence P.
; APPLICANT: KHAH, Farrah A.
; APPLICANT: RECIPIENT, Shirley A.
```

```

APPLICANT: AZIMZAI, Yalda
APPLICANT: POLICKY, Jennifer L.
APPLICANT: DING, Li
APPLICANT: GREYER, Megan
APPLICANT: ELIOT, Vicki S.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: BATRA, Sajeiv
APPLICANT: ISON, Craig H.
TITLE OF INVENTION: HUMAN KINASES
FILE REFERENCE: PI-0125 PCT
CURRENT APPLICATION NUMBER: US/10/311,034
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
60/228,056
PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0
25
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PERL Program
SEQ ID NO 5
LENGTH: 656
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID NO: 1242491CD1
US-10-311-034-5

```

```

Query Match 8.6%; Score 327.5; DB 6; Length 656;
Best Local Similarity 26.1%; Pred. No. 4.8e-17;
Matches 101; Conservative 75; Mismatches 168; Indels 43; Gaps 13;

```

```

OY 10 AMKNSF-----FROEASMLALHPCIALIGSIHPICLALPLSLNTVLSNA 63
DB 33 AKKIKNTKTSRLRLKQELVYLCCHLHPISLILANGIRPMILMELASKSGSLDRLQOK 92
OY 64 RDSFPIPLGHMLTQKIAVOIASGLAYLHKHNIIFCDLSDNLTWSDYKEHINIKLSY 123
DB 93 AS-----LTRLQHLRIALHVDAGLKYLSAMITTYDKPHNVLTLTPVNAIINKIDY 147
OY 124 GISROSEHGALGVEGTQYAPET-RPRIVYDEKVMFSYGMVLYELL-SCORPALG- 179
DB 148 GIAQCCMRGIGTSEGTGEGFAVRGIVITNOQADVSFGLILYDILTGTGRIVEGK 207
OY 180 ----BHOLOAKKSKGIRPVLGQBEVQERLOALMMECMPTREKRLALSYVSQMD 235
DB 208 FPNFDEDEIGKLPDPVAKGICAP---WPNVEKLIKQCKENQOERTSAQVEDILNS 263
OY 236 PTFATFMYELCCGKOT---AFSSQGOEYTVFMDG---KEESRNYTVVNTKEGIMEY- 287
DB 264 AELVCLTRILLPKNYVECMVATHHNSRNASIMWLGCHTBDGQSLFDLNTGEGTSEB 323
OY 268 --QRCCGEMKYSQLOVORSIM--TATEDOKIITYTLKGMCPLTPQOALTPRAYVCF 343
DB 334 ADSRLICAL---VHLPEKESMIVSGTQSLVINTEDGKRRLTEKMTDS---VCL 377
OY 344 LAVPVIR--KNSYVLAGLADGLAVF 368
DB 378 YCNFSKSKOKNPLVGTADSKLAIF 404

```

```

RESULT 4
PCT-US03-04816-56
Sequence 56, Application PC/TUS0304816
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Rosenfeld, Julie Beth
APPLICANT: Silos-Santiago, Immaculada
TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
12216, 17719, 41887, 47174, 33408, 10002, 16209, 314, 636,
27410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
TITLE OF INVENTION: 13424 MOLECULES

```

```

FILE REFERENCE: NP102-027
CURRENT APPLICATION NUMBER: PCT/US03/04816
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: US 60/360,435
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/370,121
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/373,010
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/373,908
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/377,717
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: US 60/379,949
PRIOR FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: US 60/382,409
PRIOR FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US 60/385,280
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 60/386,879
PRIOR FILING DATE: 2002-06-06
PRIOR APPLICATION NUMBER: US 60/387,536
PRIOR FILING DATE: 2002-06-10
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 56
LENGTH: 953
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-04816-56

```

```

Query Match 7.6%; Score 289.5; DB 1; Length 953;
Best Local Similarity 22.6%; Pred. No. 7.8e-14;
Matches 168; Conservative 92; Mismatches 278; Indels 207; Gaps 30;

```

```

OY 19 QEASMLHLPCHPCIALIGSIHP--LCFALPLSLNTVLSNARDSFPIPLGHMLT 76
DB 14 QEARLFGALHPNITIALGACLNPHCLTVMYAGALSRVLAARR-----VP-PHYLY 197
OY 77 OKIAVOIASGLAYLHKR--NIIFCDLSDNLTWSDYKEHIN--IKLSYGISRO 128
DB 198 -NMAVOVARGMNYLHNDAPVPIIHRDLKSNILIL--LEAIEHNHNLADYVLTDEGLARE 254
OY 129 SFHEGALVEGTPGQAEIRPRIVYDEKVMFSYGMVLYELL-SCORPALGHNOIA- 186
DB 235 WHKTKMSAAGTYAMVADPEVIRLSLEKSSDVMSFGVLLMELTGEVPRREIDALAVAG 314
OY 187 --KIKSGIRPVVLGQBEVQERLOALMMECMPTREKRLALSYVSQ 232
DB 315 VAMNLTLPPIETCEP---FAR---LLEECWDDPHGRPFGSLIKRLVIEGSALEQ 367
OY 233 MKDPTF-----ATFMYELCCGKOTAFSSQGOEYTVFMDGKEE--- 271
DB 368 MPESEFHSLOEDMKLEIOHMFDDLRTKEKELRSREBELRAAOBQFOEQLRRBOELA 427
OY 272 SNYIVVNTKEGLMEVQRCCGEMKYSQLOVORSIMTATEDOKIITYTLKG-----MCP 326
DB 428 EREMDIVEREHL-----MCOI-----SQKPRVRRKKNFRANVAK 465
OY 327 LNTPOQALDTPAVVTCFLAV---PVI--KNSYVLAGLADGLAVFVYVGRTPDSCSY 381
DB 466 LREGSSHISLPSPGFEHKITVOASPLDKRKGSDGASPPASPIPLRLAIRLTPVD--- 521
OY 382 LGSHTANNSKESIADEDARONPYVK-----AMEYVN 413
DB 522 -CGGSSSGSSSGSGSTWSRGP-PKKEELVGGKKKGRTWSPSTIQKERVGGEERLKGIG 579
OY 414 SGEVWYSNGPPL-----LVICASLEICRL-----EPYMAPSWTSYV 453
DB 580 ESKKQWSSAPNLGSPKTPPIAPGPASLNEDEFAEADGSSVPPSPYSTPSTLSVPL 639
OY 454 CSSEGRGEVWVCLDKANSLVYHSTYQL-----CA----- 486

```

```

Db      640 PAPSPGAPAPWEPPTSPAPFARWGARRRCDLALLGATLLAGVAGADVAEARAAGE 699
QY      487 ---RYCGVPSPLRDMFPVRLDTEPPASHTANPKYBEDSLADVSIMSEELGQIILI 543
Db      700 EORRWLDGLFFPRAGRP-RGL--SEPARPHGRREDVCGGLAP-----SATLV 746
QY      544 HOESLTDYCSMSYSSSPRQOARSPSLPSSPASSSSVP-----FSTDE 589
Db      747 SLSVSVDCHSTRSLRSLSDDEANAPAPSPPPAPPTPSPSTNPLVDLESEFKDPR 806
QY      590 DS-DMLHTPGAASDRSEHDLTPMDG 613
Db      807 QSLTPHTVTAACAVSRGHRRTPSDG 831

```

```

RESULT 5
US-10-369-022-56
; Sequence 56, Application US/10369022
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
; TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
; TITLE OF INVENTION: 27410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
; TITLE OF INVENTION: 95431, 22245, 2387, 16638, 35054, 16314, 1613, 1675, 9569 OR
; TITLE OF INVENTION: 13424 MOLECULES
; FILE REFERENCE: MPI02-027P1RNMNM
; CURRENT APPLICATION NUMBER: US/10/369,022
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/360,495
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/370,121
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/373,010
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/373,908
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/377,717
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 60/379,949
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/382,409
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/385,280
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/386,879
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/387,536
; PRIOR FILING DATE: 2002-06-10
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-022-56

```

```

Query Match      7.6% Score 289.5; DB 6; Length 953;
Best Local Similarity 22.6% Pred. No. 7.8e-14;
Matches 168; Conservative 92; Mismatches 278; Indels 207; Gaps 30;
QY      19 OASMLAHQHCYVALISISHP--LCFALBLAPLSINTVLSSEARSSFIPLGHMT 76
Db      144 OEARLFGALQHPNITIALRGACINPPLCLVMEYARGALSRVLAGRR-----VP-PHYLV 197
QY      77 QRIAYOISGLAYLHKR---NIIFCDKSDNIIWMSLDVKEHIN-----IKLSDGISRQ 128
Db      138 -NMAVOAGNMYLHNDAPVPIIHRDLKSINILI--LEAIEHNNLADTVLKITDFGLARE 254

```

```

QY      129 SFHEGALGVEGTPGXYOAPRIPIRYDEKYMFSYGMVLYELLSGORPALGHQLOIA-- 186
Db      255 WAKTTMSAAGTAYMAPEYIRLSLESKSSDWSFVLLMELLTGVEPYREIDALAAVAG 314
QY      187 ---KRLSGIRPYLGOPPEEYOFRLQALMECDWTPKRRP-----LALSVSQ 232
Db      315 VAMNKLTLPIPTCPPEP-----FAR---LLEECDDPPHPRDQSGILKLEVIYEGSALQ 367
QY      233 MKDPTF-----ATFTELCCGKTAFSSGQGYTYVFNQKEE--- 271
Db      368 MPLESHSLQEDWKLEIQHMFDLRTKEKELRSREBELRAAOEORFOEOLRRRQELA 427
QY      272 SRNYTVNTEKGLMEYQRMCOPKMYSCOLOQRSMTATEQKIITYLKG-----MCP 326
Db      428 ERMDLIVERLHL-----MCL-----SCKPRVRRKRRKRAVLK 465
QY      327 LNTPOALDTPPAVVTGFLAV---PVI--KNSYLVLAGIADIVAFVVRGTPKDCSY 381
Db      466 IARGSSHILSPGFEHKIYVOASPTLDKRGSGASPPASPTIIRLRIPLTPVD----- 521
QY      382 LCSHTANRSKFSIADEDANONPYVK-----CA-----AMEVYN 413
Db      522 -CGSSSSSGSSGSGTWSHGP-PKKEELVGGKKRGRTWGPSTLOKERVGEERLKG 579
QY      414 SGSEWYNSGPGI-----LVIDCASLEICRL-----EYMAPSMYTSYV 453
Db      580 ESKQKSSAPNKGSKPKTPIAPGFASLNEEFAEADGSSVPSYSTYLSVL 639
QY      454 CSSEGEVEVWCIDDKANSILVYHSTYQL-----CA----- 486
Db      640 PAPSPGAPAPWEPPTSPAPFARWGARRRCDLALLGATLLAGVAGADVAEARAAGE 699
QY      487 ---RYCGVPSPLRDMFPVRLDTEPPASHTANPKYBEDSLADVSIMSEELGQIILI 543
Db      700 EORRWLDGLFFPRAGRP-RGL--SEPARPHGRREDVCGGLAP-----SATLV 746
QY      544 HOESLTDYCSMSYSSSPRQOARSPSLPSSPASSSSVP-----FSTDE 589
Db      747 SLSVSVDCHSTRSLRSLSDDEANAPAPSPPPAPPTPSPSTNPLVDLESEFKDPR 806
QY      590 DS-DMLHTPGAASDRSEHDLTPMDG 613
Db      807 QSLTPHTVTAACAVSRGHRRTPSDG 831

```

```

RESULT 6
US-60-426-964-2
; Sequence 2, Application US/60426964
; GENERAL INFORMATION:
; APPLICANT: Sinos-Santiago, Immaculada
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: pain and painful disorders using 16314, 1613, 1675 or 9569
; FILE REFERENCE: MPI02-189PIM
; CURRENT APPLICATION NUMBER: US/60/426,964
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 953
; TYPE: PRT
; ORGANISM: homo sapiens
US-60-426-964-2

```

```

Query Match      7.6% Score 289.5; DB 7; Length 953;
Best Local Similarity 22.6% Pred. No. 7.8e-14;
Matches 168; Conservative 92; Mismatches 278; Indels 207; Gaps 30;
QY      19 OASMLAHQHCYVALISISHP--LCFALBLAPLSINTVLSSEARSSFIPLGHMT 76
Db      144 OEARLFGALQHPNITIALRGACINPPLCLVMEYARGALSRVLAGRR-----VP-PHYLV 197
QY      77 QRIAYOISGLAYLHKR---NIIFCDKSDNIIWMSLDVKEHIN-----IKLSDGISRQ 128

```

```

Db 198 -WMAVAVGAMNYLHNDAPVPIIHRDLKSINILI--LEAIENHNADLYLKITDFGLAE 254
Qy 129 SFHEGALVGEPPGYAPFIRPRIYDEKVDMSYGMVYELLSGORPALGHQLOIA-- 186
Db 255 WHTKTSAGTYAMWAPFIRLSLFSKSDVMSFVLLMELLTGEVPREIDALAVAG 314
Qy 187 ---KLSKIRIPLVAGPEVQFRRLQALMMECHDTPEKRP-----LALSVSQ 232
Db 315 VAMNKITLPISTCEP-----FAR---LLEECMDPDPHGRPDGSLTKRLEVEIQSALFQ 367
Qy 233 MKDPTF-----ATFMYELCCGKOTAFSSGCGEYTVFMDGKEE--- 271
Db 368 MPLESHSIQEDMKLEIQHMFDDLRTKEKELNSREBELRAAOEQRFQEOQLRRRQELA 427
Qy 272 SRNTYVNTKGLMEVQRCCPGMVSQLOVQRLMTATEOKIITYLKG-----MCP 326
Db 428 ERMODIVERELHLL-----MCOI-----SDEKPVRRKKNFKRAVLK 465
Qy 327 LMTPOALDTPAVVYCFILAV---PVI--KNSYVLGLADGLVAVFPVVRGTPKDCSY 381
Db 466 LREGSHISLPSGFEHKITVQASPTLDKRGSDGASPPASISIIPLRAIRLTPVD--- 521
Qy 382 LCSHTANRSKFSIADEDARQNPYVK-----AMEVYV 413
Db 522 -CGSSSSSGSGSGTWSHGR-PKKEELVGGKKKGTWGPSTLQKERVGGEERLKG 579
Qy 414 SGEVWYNSNGPGL-----LVIDCASLEICRRL-----EPYMAPSWYTSV 453
Db 580 ESKKOWSSAPNLGSKPKHTPIAPGFASLNEEFAEADGSSVPSYSPSYLSVPL 639
Qy 454 GSEGEVWYVCLDKANSLSVYHSTYQL-----CA----- 486
Db 640 LAEPSPGARAPWEPFSPAPARMWGARRRCDLALIGCATLLGAVGLADVAEARAAGE 699
Qy 487 ---RYFCVSPPLRDMFPVRLDTEPPASHTANPKVEPGSDIADVSIMYSEELGTQIL 543
Db 700 EQRRLMDLGFPPRAGRFP-RGL--SPPARPHGRREDVPGGLAP-----SATLV 746
Qy 544 HQESLTDYCSMSYSSSPRQAARSSSLPSSPASSSSV-----ESTDCE 589
Db 747 SLSSVSDCNSTRSLRSDSDENAPAPSPPPAPPTPSPSTNPLVDLELESFKDPR 806
Qy 590 DS-DMLHTPGAASDRSEHDLTPMDG 613
Db 807 GSLTPHTVTAACAVSRGHRRTPSDG 831

```

RESULT 7
US-10-263-929-187

```

; Sequence 187, Application US/10263929
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Alzheimer's Disease Linked Genes
; FILE REFERENCE: LSD-07417
; CURRENT APPLICATION NUMBER: US/10/263,929
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 187
; LENGTH: 954
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-263-929-187

```

Query Match 7.5%; Score 285; DB 6; Length 954;

Best Local Similarity 22.8%; Pred. No. 1.8e-13; IndeIs 208; Gaps 31;

Matches 170; Conservative 93; Mismatches 275; Indels 208; Gaps 31;

Qy 19 OASMLHALQPCIVALLIGISIH--LCFALDLAPLSSINTVLSSEVARDSPFIPGLHMT 76

Db 144 QCARLFGALQHNITIALRGACLNPHILGLVMEYANGALSLRYLAGRR-----VP-PHYLV 197

```

Qy 77 QKIAVQIASGLAYLHKK---NIIFDLKSDNLLWSLDVKEHIN-----IKLSDYGISRQ 128
Db 198 -WMAVAVGAMNYLHNDAPVPIIHRDLKSINILI--LEAIENHNADLYLKITDFGLAE 254
Qy 129 SFHEGALVGEPPGYAPFIRPRIYDEKVDMSYGMVYELLSGORPALGHQLOIA-- 186
Db 255 WHTKTSAGTYAMWAPFIRLSLFSKSDVMSFVLLMELLTGEVPREIDALAVAG 314
Qy 187 ---KLSKIRIPLVAGPEVQFRRLQALMMECHDTPEKRP-----LALSVSQ 232
Db 315 VAMNKITLPISTCEP-----FAR---LLEECMDPDPHGRPDGSLTKRLEVEIQSALFQ 367
Qy 233 MKDPTF-----ATFMYELCCGKOTAFSSGCGEYTVFMDGKEE--- 271
Db 368 MPLESHSIQEDMKLEIQHMFDDLRTKEKELNSREBELRAAOEQRFQEOQLRRRQELA 427
Qy 272 SRNTYVNTKGLMEVQRCCPGMVSQLOVQRLMTATEOKIITYLKG-----MC 325
Db 428 ERMODIVERELHLL-----MCOI-----SDEKPVRRKKNFKRAVLK 465
Qy 326 PLNTPOALDTPAVVYCFILAV---PVI--KNSYVLGLADGLVAVFPVVRGTPKDCS 380
Db 466 KLRGSHISLPSGFEHKITVQASPTLDKRGSDGASPPASISIIPLRAIRLTPVD--- 522
Qy 381 YLCSHTANRSKFSIADEDARQNPYVK-----AMEVY 412
Db 523 -CGSSSSSGSGSGTWSHGR-PKKEELVGGKKKGTWGPSTLQKERVGGEERLKG 579
Qy 413 NSGEVWYNSNGPGL-----LVIDCASLEICRRL-----EPYMAPSWYTSV 452
Db 580 GGSCKOWSSAPNLGSKPKHTPIAPGFASLNEEFAEADGSSVPSYSPSYLSVPL 639
Qy 453 VCSSEGEVWYVCLDKANSLSVYHSTYQL-----CA----- 486
Db 640 LAEPSPGARAPWEPFSPAPARMWGARRRCDLALIGCATLLGAVGLADVAEARAAGE 699
Qy 487 ---RYFCVSPPLRDMFPVRLDTEPPASHTANPKVEPGSDIADVSIMYSEELGTQIL 542
Db 700 EQRRLMDLGFPPRAGRFP-RGL--SPPARPHGRREDVPGGLAP-----SATLV 746
Qy 544 HQESLTDYCSMSYSSSPRQAARSSSLPSSPASSSSV-----DCEDSDMLHPR 597
Db 747 VLSVSDCNSTRSLRSDSDENAPAPSPPPAPPTPSPSTNPLVDLELESFKDPR 806
Qy 598 GAASDRSE-----HDLTPMDG 613
Db 807 GSLTPHTVTAACAVSRGHRRTPSDG 832

```

RESULT 8
US-10-263-929-193

```

; Sequence 193, Application US/10263929
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Alzheimer's Disease Linked Genes
; FILE REFERENCE: LSD-07417
; CURRENT APPLICATION NUMBER: US/10/263,929
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 193
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-263-929-193

```

Query Match 7.4%; Score 283.5; DB 6; Length 1002;

Best Local Similarity 22.8%; Pred. No. 2.5e-13; IndeIs 227; Gaps 33;

Matches 172; Conservative 99; Mismatches 257; Indels 227; Gaps 33;

Qy 9 DAMKRFSEPRQASMLHALQPCIVALLIGISIH--LCFALDLAPLSSINTVLSSEVARDSP 66

```

Db 146 DAAAESEVRREARLFAMLRHPNIIQLRGVCLRPHCLVLEFARGALNRAALAAADP 205
QY 67 SFILGHMLTOKI-----AYOIASGLAYLHK---NIIFCDLKSNDIILWMSLDVREH 115
Db 206 R-APGRARRIPPOVLVNNVAOJARGMILHBEAVVPIILHRDLKSSNILL-LEKIEH 261
QY 116 INI-----KLSDYGSROSFEHAGLVECTPGYOAPELRPIRYVDEKDMFSYGVLYEL 170
Db 262 DDIKMLKLTIDFGLARHWRHTTMSAAGTYAMMAPEVIRSSLSKSGSDIMSGLIMEL 321
QY 171 LSGORPALGHOLQIA-----KLSKGRIPVLGPEEVOFRRLQALMCEMCDTPKREKPL 225
Db 322 LTGEVPRGIDGLAVAGVANLKLPIRSTCPEP-----FAKLMCEMDPHR-- 372
QY 226 ALSVYVSKDPTFATFMEILCCGKOTAFSSOGGEYTVVFGDK-EESRNTYVNT-EKG 283
Db 373 -----PSFALLIIOQLALIEEAVLTNMPQESFHMQMDWMLKLEIQMSELTREKE 422
QY 284 LM-----EVOVMCCPGMNVSOLO-----VQSLMTATEODKITYITILKMCPLNTPQA 333
Db 423 LRSREELSRALQOKSOELLRRQOLAREIDVLERELNVLIFOLS-----QEA 474
QY 334 LDPRAVVTCLAVPVIRKNSYLVAG---LADGLVAVP-----VYRGTP-----KDS 378
Db 475 -----PHYKRRKGRFRGRRLKLDGHRISLPSDFOKITVOASPTLDKRRSS 521
QY 379 CSYLCSHTAN-----RSKFSIADSD-----ARQNPYKAMEVYVNSGSEVYYSN 422
Db 522 DSGICSPGSPMLRPLRAIOLTSDENNKTRGRNNVFRQEDFE---DVKRSFKKCCWTW 577
QY 423 GPGLLVIDCASLEICRLEPYMAPSMTSYVCSGEGEEVYVWCLD---KANLWYHS 479
Db 578 GP-----SSVQTKERPEGRERVRPLSGNSPWSLSLIRKSO 612
QY 480 TTYOLCARVCGVSPRLRDEPV-----RPLDTEPPAASHRTANPVY----- 520
Db 613 KTRPL-ASLEVPDQPGCEQKLVPEGLEHRRKPKTKRPGQAHVGLPLCKXQSRDESEAE 671
QY 521 -----PEGSIADVS-----IMYSEIGTOI-LIH----- 544
Db 672 SREBGSKSGSPVNNVGAAPMLRKKTESALCEGMLASMAGLDVRKLGQAAPAKPAPKM 731
QY 545 -----QESLDTYCSMSYSVSSPPRO--NARSPSS-----LPSPPASSSVFSTDC-- 588
Db 732 EKKEEGALOPASRCQSSPSSILQPSAGRAPSGGSTILLPSABSHSKSSLSMKCLDQAG 791
QY 589 -EDSDMLHT-----PGAASDRSEHDITP 610
Db 792 KEESISIGNARDLCGPTTLFPDPGSAAPESGCEIIP 826

```

```

RESULT 9
PCT-US02-39126-13
; Sequence 13, Application PC/TUS0239126
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: KABLE, Amy E.
; APPLICANT: CHIEF, David
; APPLICANT: WILSON, Amy D.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: GORVAD, Ann E.
; APPLICANT: HAFALIA, April J. A.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: RAMKUMAR, Jayalakshi
; APPLICANT: JIN, Pei
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: LEHR-MASON, Patricia M.
; APPLICANT: KHARE, Reena
; APPLICANT: LEE, Sally
; APPLICANT: HAWKINS, Phillip R.

```

```

; APPLICANT: BECHA, Shanya D.
; APPLICANT: LEE, Soo Yeun
; APPLICANT: SPRAGUE, William W.
; APPLICANT: ZEBARADIAN, Yeganeh
; TITLE OF INVENTION: KINASES AND PHOSPHATASES
; FILE REFERENCE: PF-1315 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/39126
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/340,235
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/343,007
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/343,546
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/354,388
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/357,675
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 3810039CD1
PCT-US02-39126-13

```

```

Query Match 7.28; Score 273.5; DB 1; Length 1081;
Best Local Similarity 31.38; Pred. No. 1.7e-12;
Matches 78; Conservative 41; Mismatches 93; Indels 37; Gaps 8;

```

```

QY 3 RHLRATPMKNFSEFREASMLALHCPICVALLIGSIHP--LCFALSLAPLSLNTVLS 60
Db 174 RHPDEDISOTIENVROEALFLMLKHPNIIQLRGVCLKEPNCLVNEFARGPPLNRYLS 233
QY 61 ENARDSSEFPLGHMLTOKI-----AYOIASGLAYLHK---NIIFCDLKSNDIILWMSL-- 110
Db 234 -----GRIPIIDLNNVAOJARGMANTLHBEAVVPIILHRDLKSSNILLQKVEN 282
QY 111 -DYKHEHNILSDYGSROSFEHAGLVECTPGYOAPELRPIRYVDEKDMFSYGVLYEL 169
Db 283 GDSINMT-LKITDFGLARHWRHTTMSAAGTYAMMAPEVIRASMTFSKSGSPWSYGVLYME 341
QY 170 LSGORPALGHOLQIA-----KLSKGRIPVLGPEEVOFRRLQALMCEMCDTPKREKRP 224
Db 342 LTGEVPRGIDGLAVAGVANLKLPIRSTCPEP-----FAKLMCEMDPHSRP 394
QY 225 LALSVYVSKM 233
Db 395 SFTNIIDQL 403

```

```

RESULT 10
US-10-210-130-130
; Sequence 130, Application US/10210130
; GENERAL INFORMATION:
; APPLICANT: zerhusen, Bryan D.
; APPLICANT: Paturajan, Meera
; APPLICANT: Kexuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: LI, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glenda

```

APPLICANT: Ji, Weizhen
APPLICANT: Gorman, Linda
APPLICANT: Vernet, Corline A.M.
APPLICANT: Leite, Mario W.
APPLICANT: Guo, Xiaojia Sasha
APPLICANT: Anderson, David W.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Gerlach, Valerie
APPLICANT: Burgess, Catherine E.
APPLICANT: Khrantsov, Nikolai V.
APPLICANT: Ort, Tatiana
APPLICANT: Ellerman, Karen
APPLICANT: Rastelli, Luca
APPLICANT: Agee, Michele L.
APPLICANT: Chaudhuri, Amitabha
APPLICANT: Chant, John S.
APPLICANT: DiPippo, Vincent A.
APPLICANT: Edinger, Shlomit R.
APPLICANT: Eisen, Andrew J.
APPLICANT: Gangoli, Esha A.
APPLICANT: Gioc, Loic
APPLICANT: Ooi, Chean Eng
APPLICANT: Rothenberg, Mark E.
APPLICANT: Spaderna, Steven K.
APPLICANT: Hult, Tord
APPLICANT: Liu, Xiaohong
APPLICANT: Taupier, Raymond J., Jr.
APPLICANT: Catterton, Elina
APPLICANT: Shenoy, Suresh G.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-416C (Cura-716 SMT)
CURRENT APPLICATION NUMBER: US/10/210,130
CURRENT FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: 60/309,501
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 60/316,508
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 60/354,655
PRIOR FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: 60/310,291
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 60/383,887
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: 60/310,951
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: 60/323,936
PRIOR FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: 60/381,039
PRIOR FILING DATE: 2002-05-16
PRIOR APPLICATION NUMBER: 60/311,292
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/311,979
PRIOR FILING DATE: 2001-08-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 369
SOFTWARE: CuroSeq1st version 0.1
SEQ ID NO 130
LENGTH: 1096
TYPE: PRT
ORGANISM: Homo sapiens
US-10-210-130-130

Query Match 7.2%, Score 273.5; DB 6; Length 1096;
Best Local Similarity 31.3%, Pred. No. 1.7e-12;
Matches 78; Conservative 41; Mismatches 93; Indels 37; Gaps 8;

Qy 3 RLRLRTDAKMFSEFQESMHLQPCYALIGISHP--LCFALSLAPISLNTVLS 60
Db 152 RHDPEDISQTEIENVRQEKLFAMLKHPNIIALRGVCKEPLCLVMEFARGGPLNRYLS 211

Qy 61 ENARSSSTIPGLHMLTKR---AVQIASGLAYLKR---NIIPFDLSNDLTVMSL--- 110
Db 212 -----GKRIPDPLVMAVAQIAAGMYLHDEAIVPIIHRDLSSNIIILQKVEN 260

Qy 111 -DYKEHINIKLSDYGISRQSFHEGALVECTPGYQAPRPRIVYDERKVMFSYGVLYE 169
Db 261 GDLNKR-LKITDGLAREHRTTKMSACTYAMNAPEVIRASFSKSGSVWYGVLYE 319

Qy 170 LLSGQRPALGHQLOIA-----KLSKGRPVLSQPEEVQFRRLQALMECWDTPKPKRP 224
Db 320 LITGEVFRGIDGLAVYGVAMNKLALPISTCEP-----FAKLMECDWNPDPHSRP 372

Qy 225 LALSVSQM 233
Db 373 SFNIIIDPL 381

RESULT 11
US-10-052-648A-42
Sequence 42, Application US/10052648A
GENERAL INFORMATION:
APPLICANT: Anderson, David
APPLICANT: Burgess, Catherine
APPLICANT: Casman, Stacie
APPLICANT: Colman, Steven
APPLICANT: Edinger, Shlomit R.
APPLICANT: Ellerman, Karen
APPLICANT: Gerlach, Valerie
APPLICANT: Gunther, Erik
APPLICANT: Kerkuta, Ramesh
APPLICANT: MacDougall, John R.
APPLICANT: Mehraban, Fuad
APPLICANT: Patturajan, Meera
APPLICANT: Rothenberg, Mark
APPLICANT: Shinkets, Richard
APPLICANT: Smltson, Glenda
APPLICANT: Spytek, Kimberly A.
APPLICANT: Stone, David J.
APPLICANT: Vernet, Corline A.M.
APPLICANT: Zerhusen, Bryan D.
TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
FILE REFERENCE: 21402-250 (Cura-550)
CURRENT APPLICATION NUMBER: US/10/052,648A
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: 60/262,454
PRIOR FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: 60/272,920
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/284,549
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/303,229
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,605
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/269,098
PRIOR FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: 60/264,159
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/271,855
PRIOR FILING DATE: 2001-02-27
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 97
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 42
LENGTH: 251
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: pkinase,
OTHER INFORMATION: Protein kinase domain sequence
US-10-052-648A-42


```

: PRIOR FILING DATE: 2001-02-13
: PRIOR APPLICATION NUMBER: 60/226,703
: PRIOR FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/330,293
: PRIOR FILING DATE: 2001-10-18
: PRIOR APPLICATION NUMBER: 60/332,127
: PRIOR FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: 60/280,859
: PRIOR FILING DATE: 2001-04-02
: PRIOR APPLICATION NUMBER: 60/310,797
: PRIOR FILING DATE: 2001-08-08
: PRIOR APPLICATION NUMBER: 60/268,646
: PRIOR FILING DATE: 2001-02-14
: Remaining Prior Application data removed - See File Wrapper or PALM
: NUMBER OF SEQ ID NOS: 547
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 151
: LENGTH: 254
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: pkinase,
: OS-10-074-978A-151

```

```

Query          2 LRHRRAIDAMNTEFEPQOASMLHALGHCVALIGI--SIHPCFALEAPLSIMTVL 59
Db             29 IKIKAKSLSSKKRRFLREIOILRRLSHPNIVRLGVFEEDDHLYLVMEYEGSDLPDYL 88
               :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Query         60 SENARDSSEFPDGLMLOTOKIAVOIASGLAYLAKNNIIFCDEKSDNIITWSSLDVKEHNINX 119
Db            89 RRNG-----LLTSEKEAKKIALQILRGLEYHSNGIYVRDLKPENIL-----LDENGTVK 138
               :|||:|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Query        120 LSDYGISRO---SHEGALSVGECPGOAPHIRRIYIDEVDYMFSGMYIELLSQRP 176
Db           139 IADGLKRLKESSSYEKLTFVTGPEYMAPREVLEGRTYSRSKPDVWSLGVIITYELLTSKLP 198
               :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Query        177 ALGHHOAIKAKKISKGRIPVLGP-----EEVOFRRLQAILMEBCWPTKPEKRPLATSY 230
Db           199 FPGIDPLEELFRINE--RPRLRPLPENCSEE-----LKQDIKKCKLNKDDEKRTANEIL 251
               :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Query         231 S 231
Db            252 N 252

RESULT 13
US-10-072-012-800
; Sequence 800, Application US/10072012
GENERAL INFORMATION:
APPLICANT: Tchernerov, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Ganggoli, Bsha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R. A.
APPLICANT: Pena, Carol E. A.
APPLICANT: Furtek, Katarzyna
APPLICANT: Groose, William M

```

7

Query Match 7.1%; Score 271; DB 6; Length 256;
Best Local Similarity 32.4%; Pred. No. 3.5e-13;
Matches 78; Conservative 42; Mismatches 93; Indels 28; Gaps

```
OY      2 LHRHATDAMKNSEFRQDSALHALQHPCIALIGI--SIHPLCFALSLAPSSINTVU    59  
          ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db       29 IKILKRSLSEKKRKFRLREIQILRRISHPNIVLLGVFEEDHLXLYMVEYMEGGDFDYL   88  
  
OY      60 SENADSSFIPGNHLJTOKIAVOIASGLAYLKNNIFPDCLKSDNLTWMSLDVKHEHNK    119  
          :|||:|||:|||:|||:|||:|||:|||::~|||::|  
Db       89 RRNG-----LTLISEKEAKTIALQLIRGLEYLSRGIVHDLPENFL----LDENGTVK   138  
  
OY     120 LSDYGISRK---SFHGAGLVGEPTFGPYCAPELRPRIVDKYDMESYGAUYELLSGQP   176  
          ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     139 IADFELARKLESSYEKLITTEVGTPDEVMAPEVLDEGRGSYSKVDMSELGIYLVELLNKTPL   198  
  
OY     177 ALGHQLOIARUKLSGIRPVLGCP-----EEVQPRRDOLAMECHDTLPKPRLPALSXY   230  
          ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     199 FPGIDPLELELFRIKE--REPRLPLPPNCSEE---LKDLIKCKINPKDPKRPTRAKEILL 251  
  
OY           231 S 231  
Db         252 N 252
```

RESULT 14
US-10-072-012-857
Sequence 857, Application US/10072012
GENERAL INFORMATION:
APPPLICANT: Tchernev, Velizar
APPPLICANT: Spytek, Kimberly
APPPLICANT: Zernusen, Bryan
APPPLICANT: Patturajan, Meera

```

APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupler Jr, Raymond J.
APPLICANT: Gusev, Vladimr Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosee, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072.012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 857
LENGTH: 256
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Protein kinase
US-10-072-012-857
Query Match 7.1%, Score 271; DB 6; Length 256;
Best Local Similarity 32.4%; Pred. No. 3.5e-13;
Matches 78; Conservative 42; Mismatches 93; Indels 28; Gaps 7;
OY 2 LRHRLATAMKNSFROBASMALHLPCHYALLGI--SHPLCFALSLALSSLYNL 59
Db 29 IKILKRSLSEKKKFLREIOTLRSLSHNIVRLGVFEEDHLVLMYMEGGDLFDYL 88
OY 60 SENARDSFIRPLGHLTQKIAYQIASGLAYLHKNNIIFCDLSNDNLIWLSLVKKEHINK 119
Db 89 RRNG-----LLISEAKKIALQILRGLEYLHSRGIVHNDLPENL-----LDENGTVK 138
OY 120 LSDYGISRQ---SFIEGALGVEGTPGYOAPELIRPIVYDEKVDMSYGVNLYELLSGQRP 176
Db 139 IADFLAKRLESSYEKLTTFVGTPEYMAPEVLGEGRGYSKADVMVMSGLVILYELLTKLP 198
OY 177 ALGHQDQIAKKLSGIRPVLAGP-----EEVQRRRLQALMMECMDTPKPRPLALSV 230

```

```

: Sequence 5 Application US/09977261
: Publication No. US20030054527A1
: GENERAL INFORMATION:
: APPLICANT: ULLRICH, AXEL
: APPLICANT: GISHIKY, MIKHAEL
: APPLICANT: SURES, IRMINGARD
: TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
: FILE REFERENCE: 038602/1259
: CURRENT APPLICATION NUMBER: US/09/977,261
: CURRENT FILING DATE: 2001-10-16
: PRIOR APPLICATION NUMBER: 08/732,545
: PRIOR FILING DATE: 1994-04-22
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 5
: LENGTH: 2770
: TYPE: DNA
: ORGANISM: Unknown Organism
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (366)..(1880)
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
US-09-977-261-5

Alignment Scores:
Pred. No.: 2,03e-17 Length: 2770
Score: 247.00 Matches: 99
Percent Similarity: 44.32% Conservative: 65
Best Local Similarity: 26.76% Mismatches: 129
Query Match: 6.46% Indels: 78
DB: Gaps: 15

US-09-836-392-21 (1-728) x US-09-977-261-5 (1-2770).
Qy 15 SerGluPheArgGngInuAlaSerMetLeuHAlaLeuGlnHisProCysIleValAla 34
Db 1179 AATGACTCTCCGAGGAGGCGACAGATTAATAAACCCTAAGACATCCAAAGCTTATCCG 1233
Qy 35 LeuIleGlyIleSerIle-----HisProLeuCysPheAlaLeuGlnLeuAlaProLeu 52
Db 1239 CTTATGCTGTTTGCACCTTAGAAGATCCAAATTATATTATTATACAGAGTTGATGAGACAT 1298
Qy 53 SerSerLeuAsnThrValLeuSerGluAsnAlaIArgAspSerSerPheIleProLeuGly 72
Db 1299 GGAAGTCGCGAAGATATCTC---CAAAATGACACTGATCAAAATC----- 1343
Qy 73 HisMetLeuThrGlnIleSile-----AlaTyrGlnIleAlaSerGlyLeuAlaTyrLeu 90
Db 1344 ---CATGTGACTCAACAGGTAAGATGACATGGCGGACACAGATGGCTCTGGAAATGGCTATCTG 1400
Qy 91 HisIysIysAsnIleIlePheCysAspLeuIlySerAspAsnIleLeuValItrSerIleu 110
Db 1401 GAGTCTCGGAACCTACATTCACAGAGATCTGGCGCCAGAAATGCTCTC----- 1448
Qy 111 AspValIysGluHisIleAsnIleIysLeuSerAspTyrIleTyrIleSerArg----- 127
Db 1449 ---GTTGGTGAACATATATATCTCAAAAGTGCAGATTTTGGACTTCCAGAGTTTTTAAG 1508
Qy 128 -----GlnSerPheHisGlnGlyAlaLeuGlyIleAlaGlnGlyThr 140
Db 1506 GTAGATTAAGAAGCATCTATGAATCTAGACACGGAATAAATAGCTGCCGGTGAAGTGGACT 1565
Qy 141 ProGlyTyrGlnAlaProGlnIleArgProArgIleValIlyrAspIleIysValAspMet 160
Db 1566 -----GGCGCCGAAGCCATTCGTAAGTAATAATTCGATTAAGTCCAGTGA 1611
Qy 161 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSer---GlyGlnArgProAlaLeuGly 179
Db 1614 TGCGTCAATTGGAAACCTCTCTTAAAGAAATCATTAAGTGAAGCAAAATGGCTTACAGTGGT 1672
Qy 180 HisHisGlnLeuGlnIleAlaIlyLeuLeuSerIlyGlyIleArgProValLeuGln 199

```



```

Db 2539 CCGACCGACGCCCGCTCCCGCTATGCTTCGCTCAAGACGCC----- 2590
Oy 600 aserapserglnhisaspleuthrproketaapgly 613
Db 2591 ----GACTC-CCCGCCCACTCTCTGCACCCCTGTCTGTA 2625

RESULT 12
US-09-938-842A-903
; Sequence 903, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 903
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-903

Alignment Scores:
Pred. No.: 3,41e-19 Length: 1638
Score: 259.50 Matches: 68
Percent Similarity: 48.88% Conservative: 41
Best Local Similarity: 30.49% Mismatches: 81
Query Match: 6.79% Indels: 33
DB: 9 Gaps: 8

US-09-836-392-21 (1-728) x US-09-938-842A-903 (1-1638)
Oy 16 Gluphearglnglnalasermetleuthrhisalaleuglnhisprocyalevalaleu 35
Db 955 GAGTTGGCCCAAGAGCTTATTGTGAGGAAGTTACACAAAAAGTTGTTAGTTC 1014
Oy 36 lilegyliserlehispro-----leucysphealaleuglnleualaproleuser 53
Db 1015 ATTGTGCTTCGACCAACGCTCCACATCTGTATCTTACAGATTCATGCCCGTGA 1074
Oy 54 Serleuanthvalleusergluasnalargaspersepherleproleuglnhis 73
Db 1075 AGTGTATGACTATCTACAC-----AAGCAAAAGGCGCTTTAAGCTTCAACT 1125
Oy 74 Metleuthrglnlysilealatyrglnlealaserglyleualatyrlenthislylys 93
Db 1126 TTGTT-----AAGTAGCTATAGATATTGGCAAGGAGGAGCTACTTACCAAAAT 1179
Oy 94 Asnillelepiecyaspleulysersaspsanilleuvaltyrserleuaspvallys 113
Db 1180 AACCAATATTCACAGAGATTGAGCTGCCAACCTCTTA-----ATGGAC 1224
Oy 114 GluHsIleAanIlelyleuseraspyrgylieserarglnserphenisgluly 133
Db 1225 GAAATATGAGGTGTTAAGGTTGCAGACTTGGGGTGGCTAGGTAAGCAACAACTGGA 1284
Oy 134 Alaleuglyvalglu---glythrproglytyrglnalaproglnleargproargile 152
Db 1285 GTTATGACAGCTGAACAGCAATATGCTGGAGTGGCTCCAGAGCTATAGAAACAG 1344
Oy 153 Valtyrappglulysvalaspmetheserlyrglymetvalleuvalleuvalleu 172

```

```

Db 1345 CCANATGATCAAGAGCTGACGATTCAGCTACGGGATTTGCTATGGAGTTGTA 1404
Oy 173 Glylnarproalaleuglyhisleaglenthleuglnlealalelylsleuserly 192
Db 1405 GGG-----AAGGA 1413
Oy 193 lileargprovalleuglylnprogluvalglinpheargyleuglnalaleu 212
Db 1414 TTAGGCCCAACAAT-----CCAAAGAACGCGAT-----CCGAATTTGGCAGGCTAT 1464
Oy 213 Metglucystrpaspthrlylserproglulysarproleualaleuserval 232
Db 1465 GAGAGATTGTGGAGCATGATTCGACGACAGACGACTTCTCAGAGATCATAGACG 1524
Oy 233 Metlyasp 235
Db 1525 CTTCAGAG 1533

RESULT 13
US-09-938-842A-1577
; Sequence 1577, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1577
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1577

Alignment Scores:
Pred. No.: 5,01e-18 Length: 2211
Score: 251.00 Matches: 67
Percent Similarity: 50.21% Conservative: 53
Best Local Similarity: 28.03% Mismatches: 73
Query Match: 6.57% Indels: 46
DB: 9 Gaps: 9

US-09-836-392-21 (1-728) x US-09-938-842A-1577 (1-2211)
Oy 15 Sergluphearglnglnalasermetleuthrhisalaleuglnhisprocyaleval 34
Db 1522 ACGAGTGCACAAAGAGATCAACTTTGTGAAGAACTGACATCCCAATGTGCTACTA 1581
Oy 35 leuilegyliserlehisproleucys-----phealaleuglu 48
Db 1582 TTTATGGGA-----GCAGTATGTACAGAAAGAAATGCGCATATATGAGAA 1629
Oy 49 leualaproleuserleuanthvalleusergluasnalargaspersepher 68
Db 1630 TATATGCCAAGAGGAGCTTCTCAAAATCTTCAATATACAAATCAG----- 1677
Oy 69 lileproleuglnhismetleuthrglnlystlealatyrglnlealaserglyleu 88
Db 1678 ---CCATTGGACACAAAGCCCGTTTAAAGATGCGCTTGATGTTCAGGGGAATGAAT 1734
Oy 89 TyrleuHslylslyasn-----llellephecyaspleulysersaspsanilleu 106

```

ORGANISM: Homo sapien
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (482)...(3023)
 US-10-143-133-1

Alignment Scores:

Pred. No.:	9-1e-19	Length:	3558
Score:	260.50	Matches:	151
Percent Similarity:	35.76%	Conservative:	90
Best Local Similarity:	22.40%	Mismatches:	236
Query Match:	6.82%	Indels:	199
DB:	9	Gaps:	29

US-09-836-392-21 (1-728) x US-10-143-133-1 (1-3558)

```

QY 18 ArgGlnGlnAlaSerMetLeuHisAlaLeuGlnHisProCysAlaValAlaLeuIleGly 37
DB 965 CGCCAGAGAGCGCGGCTCTTCCCATGCTGGCACAACCCCAACATCATTCCTCAAGGCT 1024
QY 38 IleSerIleHisPro-----LeucysPheAlaLeuGlnLeuAlaProLeuSerSerLeu 55
DB 1035 GTGTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084
QY 56 AsnThrValLeuSerGlnAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeu 75
DB 1085 AGCCAGAGCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1126
QY 76 ThrGlnIleValIleAlaTyroIleAlaSerGlyLeuAlaTyroIleHisIleVal 93
DB 1127 GTC---AACTGGGCTGTGGAGATGTCCTGGGAGATGCTACCTGACCTGCGAGGCGCTG 1183
QY 94 ---AsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValTrp-----Ser 109
DB 1184 GTGCCGATATCCACCGGATATCTCAAGTCAACACATTTGGCTGTGCAAGCCATTGAG 1243
QY 110 LeuAspValIleGlnHisIleAsnIleLysLeuSerAspTyroIleIleSerArgGlnSer 129
DB 1244 ACTGACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1303
QY 130 PheHisGlnGlyValAlaLeuGlnIleValGlnIleValTrpProGlyIleAlaProGlnIleArg 149
DB 1304 CACAAACACACAAATGAGTGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1363
QY 150 ProAlaGlyIleValIleArgProAlaValAlaSerMetPheSerTyroIleValIleValTrp 169
DB 1364 AAGGCTCCACCTTCTCTAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1423
QY 170 LeuLeuSerGlyGlnArgProAlaLeuGlnHisIleGlnIleAla-----186
DB 1424 CTGCTGACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1483
QY 187 -----LysAlaLeuSerLysGlyIleArgProValLeuGlnIleValGlnIleValGln 204
DB 1484 GCTGTTAAACAGCTCACACTGCCCATTCACCTGACCTGACCTGACCTGACCTGACCTGAC 1531
QY 205 PheArgArgLeuGlnAlaLeuMetMetGlyCysTrpAspThrIleProGlyIleValArgPro 224
DB 1532 -----TTGCGACAGCTTATGCGGAGCTGCTGGCGGAGAGAGAGAGAGAGAGAGAGAG 1582
QY 225 -----LeuAlaLeuSerValIleSerGlnMet 233
DB 1583 GACTTGCCCTCATCTCTGACAGAGTGGAGGCGCTGGAGGACACAGTCTCTCAAGGAAATG 1642
QY 234 LysAspProThrPheAlaThrPhe-----MetTyr 243
DB 1643 CGCGGAGAGCTCTTCATATTCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1702
QY 244 GluLeuCysCysGlyGlnIleValIlePhePheSerSerGlnIleGlnIleValIleVal 263
DB 1703 GAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1759
QY 264 ValPheTrpAspGlyLysGlnIleValIleValIleValIleValIleValIleValIleVal 278
  
```

```

DB 1760 GCGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1819
QY 279 AsnThrGlnLysGlyLeuMetLeuValGlnArg-MetCysCysProGlyMetLysValSer 298
DB 1820 CAGTGGAGAGCTAGAGAGTGTTCAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1862
QY 298 rCysGlnLeuGlnValGlnArgSerLeuThrPheAlaThrGlnAspGlnIleTyrIle 318
DB 1863 -TGC-----AGCAGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1889
QY 318 eTyrThrLeuLysGlyMetCysProLeuAsnThrProGlnIleAlaLeuAspThrProAl 338
DB 1889 -----1889
QY 338 aValValThrCysPheLeuAlaValProValIleLysIleAsnSerTyrLeuValLeuAl 358
DB 1890 -----ACGTGCGCGCGCGCGCGCGG-ACATTCAAGCGCACAGAGCTCCGCGCGGCA 1941
QY 358 aGlyLeuAlaAspGlyLeuValAlaValPheProVal-----Val 371
DB 1942 CGGC-----GGGAGCGGTATGACATGCCACTGCACTTCAAGACACCGATCACCGT 1992
QY 371 lArgGlyThrProLysAspSerCysSerTyrLeuCysSerHisThrAlaAsnArgSerLys 391
DB 1993 GCAGGCGCTCACCC-----GGCTTGCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2028
QY 391 sPheSerIleAlaAspGlnAspAlaArgGlnAsnPro---TyrProValIleAlaMetGln 410
DB 2029 CTTGAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2088
QY 410 uValValAsnSerGlySerGlnValTrpTyrSerAsnGlnIleProGlyLeuLeuValIleAs 430
DB 2089 GCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2128
QY 430 pCysAlaSerLeuGlnIleCysArgArgLeuGlnIleProTyrMetAlaProSerMetValTrp 450
DB 2129 -----CGACGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2143
QY 450 rSerValValCysSerSerGlnGly-----458
DB 2144 -----TCAAGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2190
QY 459 -----ArgGlyGlnIleValValTrp 465
DB 2191 CCCCAAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2250
QY 465 pCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThrThrTyrGlnLeuGly 485
DB 2251 GTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2310
QY 485 sAlaArgTyrPheCysGlyValProSerPro---LeuAlaArgMetPheProValAlaArg 504
DB 2311 TGGT-----AACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2358
QY 504 oleuAspThrGlnProProAlaAlaSerHisThrAlaAsnProLysValProGlnIleLys 524
DB 2359 TGTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2412
QY 524 pSerIleAlaAspValSerIleMetTyrSerGlnIleGlnIleValIleValIleValIleVal 544
DB 2413 GCTGTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2464
QY 544 sGlnIleSerLeuThrAspTyrCysSerMetSerTyrSerSerSerProProArgGln 564
DB 2465 -----CAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2478
QY 564 nAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSerValProPh 584
DB 2479 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2538
QY 584 eSerThrAspCysGlnAspSerAspMet-----LeuHisThrProGlyAlaAla 600
  
```

```

||||| :|||||
Db 2196 GAA-----GAGCCAAACTTCCCTCGAT-----GATTAGGA 2228
QY 458 GYATGGLYGLUVALVALTRP-CysteulaspplysalaAsnSerLeuValMet 477
Db 2229 CACAGAAAACAAATAAATTCCTTACTAGCGGCTTACTGATCTCTGGG 2288
QY 477 HISSerThrThrGlnLeuGlyAlaArgTyrPheCysGlyValProSerProLeu 497
Db 2289 AAGATGCTCAGAGAGATCTCGAGAACTGAAAGCTGGAGAGGACACCTCGG 2348
QY 497 GASpMetPheProValArgProLeu-Asp-ThrGluProProAlaAsnSerThrAla 516
Db 2349 A---ATGCTGCCACAGCTCTCCTGATGATGACTCTAGATATGCTGATGATCCCC 2405
QY 517 AsnProLysValProGluGlyAspSerIleAlaAspValSerIleMetYrSerGluGln 536
Db 2406 CAGAGAAAACAAAGAGAGTCTGATGGGTGACACCTCTCTGATCGGTGGCT 2465
QY 537 LeuGlyThrGlnIle-LeuIleHis----- 544
Db 2466 CTGGGACTGCTGACAGAACTTCATAAGCACAGCTGTGAAGAACCTGGCCCAAG 2525
QY 545 -----GlnGluSerLeuThrAspTyrCysSerMetSerSerYrSerSer 559
Db 2526 GAAGAGAGAGAAACGAGAGGGAATCTTCCAGCGGGCTTCCAGAGAGAGTCC 2585
QY 560 SerProProArgGln-----AlaAlaArgSerProSerSerLeuPro 573
Db 2586 ATCTCCCTCCCAAGCTGTCATCCACCKTGGGGAGGAGCCAGACCCCTCCGCGCA 2645
QY 574 SerSerProAlaSerSer-----SerSerValProPheSerThrAspCysGln 589
Db 2646 CTGTCAAGTCCCTGGGCGATCCTCTCCACACCTTCTTCTCCAAAGTCCGTGCGAG 2705
QY 590 -----AspSerAspMetLeu 594
Db 2706 AAGGACAGTGAAGATCCACTGTGTGACAGTGCACCTGTGACTGTGAGATGCTC 2765
QY 595 -----HisThrProGlyAlaAlaSerAspArgSerGlnHisAspLeuThrPro 610
Db 2766 ACTCCGAGATTTTGTTCCTCCACTGCCCCAGAAATGTGTCGAGCCAGCCCTCATGCCA 2822

RESULT 10
US-09-938-842A-882
; Sequence 882, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME. AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 882
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-882

Alignment Scores: 3 01e-20 Length: 1428
Score: 268.00 Matches: 66

```

```

Percent Similarity: 55.11% Conservative: 58
Best Local Similarity: 29.33% Mismatches: 79
Query Match: 7.01% Indels: 22
DB: 9 Gaps: 9

US-09-836-392-21 (1-728) x US-09-938-842A-882 (1-1428)
QY 16 GluPheArgGlnGluValSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeu 35
Db 622 GAGTTTATTCGTAAGGACACCTTTTATTCGTAAGGACACCAAAATGCTTAAGTTT 681
QY 36 IlegIlySerIleHisProLeuGlyPheAlaLeuGlnLeuAlaProLeuSerSerLeu 55
Db 662 GTTGAAGGAAATCTGAAAC--TGTATCATCAGAGATGATGACCTGAGGGCTTTTA 738
QY 56 AsnThrValIleSerGlnAsnAlaArgAspSerPheIleProLeuGlyHisMetLeu 75
Db 739 AGATCATATCTGCACAACTGACGAGAAATCC-----CTTCCTTGGACAGCTATTC 792
QY 76 ThrGlnLysIleAlaTyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysAsnIle 95
Db 793 -----GATTTTGTCTGTGATATTCCTAAAGAAATGAAATATATCTCAAGAGATA 846
QY 96 IlePheCysAspLeuLysSerAspAsnIleLeuValIlePheLeuAspValLysGlnHis 115
Db 847 GTTCATCAGAGATCTAACCAGAAACCGTGTG-----ATGACAAATGAC 891
QY 116 IleAsnIleLysLeuSerAspTyrGlyIleSerArgGlnSerPheHisGluValAlaLeu 135
Db 892 TTTCCTTGAAATGTCGATCTTGGCATAGCGTCGAGAGAGATATCTGATGTTTG 951
QY 136 GlyValGlu--GlyThrProGlyLysGlnAlaProGluIleArgProArgIleValTyr 154
Db 952 GGGGATACATGAACTTAATAGTGGATGGACCTGAAGTTTAAACGATACCAT 1011
QY 155 AspGluLysValAspMetPheSerTyrGlyMetValLeuTyrGlnLeuLeuSerIlyGln 174
Db 1012 GAGCGAAGTGGATGTTATATGTTTGGACTTCTTATATGGAAATGATGACTGGA--- 1068
QY 175 ArgProAlaLeuGlyHisHisGlnLeuGlnIleAlaLysLysLeuSer----- 190
Db 1069 -----GCATCTCCATATGAGAGAGATGAATTTGCTGAACAAATTCCTACGACTTATA 1122
QY 191 ---LysGlyLysLeuArgProValLeuGlnProGluGlnValGlnPheArgArgLeuGln 209
Db 1123 TACAGAAATATAGGCCAGTATA-----CCGACGGAATGTGCCA---GCGGCATGAAA 1173
QY 210 AlaLeuMetMetGluCysTrpAspThrLysProGluLysArgProLeuAlaLeuSerVal 229
Db 1174 GAGCTGATCGAGCGATGTTGTCTATCGCAAAACAGACAAAGACGGAATTCGTGCGAGATT 1233
QY 230 ValSerGlnMetLys 234
Db 1234 GTCAAGTCTTGAA 1248

RESULT 11
US-10-143-133-1
; Sequence 1, Application US/10143133
; Publication No. US20020197658A1
; GENERAL INFORMATION:
; APPLICANT: Yoganathan, Thillainathan
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: Cancer Associated Protein Kinase, and Its Use
; FILE REFERENCE: KINE-023
; CURRENT APPLICATION NUMBER: US/10/143,133
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,555
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3558
; TYPE: DNA

```

QY 595 -----HisThrProGlyAlaAlaSerAspArgSerGluHisAspLeuThrPro 610
 DB 2503 ACTCCGGATTGTTGTCCTCCACGTGCCCGAGAGAGTGCTGTGAGCCACCTCATGCCA 2559
 RESULT 9
 US-10-014-882-3
 ; Sequence 3, Application US/10014882
 ; Patent No. US20020107384A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Kieke, James
 ; APPLICANT: Donoho, Gregory
 ; TITLE OF INVENTION: No. US20020107384A1el Human Kinase and Polynucleotides Encoding
 ; FILE REFERENCE: LEX-0279-USA
 ; CURRENT APPLICATION NUMBER: US/102014,882
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: US 60/254,744
 ; PRIOR FILING DATE: 2000-12-11
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 3518
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-014-882-3
 Alignment Scores:
 Pred. No.: 8,64e-20 Length: 3518
 Score: 269.50 Matches: 177
 Percent Similarity: 36.66% Conservative: 102
 Best Local Similarity: 23.26% Mismatches: 228
 Query Match: 7.05% Indels: 224
 DB: 12 Gaps: 36
 US-09-836-392-21 (1-728) x US-10-014-882-3 (1-3518)
 QY 9 AspAlaMetLysAsnPheserGluPheArgGlnGluAlaSerMetLeuHisAlaLeuGln 28
 DB 741 GAGCGCGCGCGCGCGCGCGAGAGCGTGGCGCGCGCGCGCTGCTTCGCCATGCTGGC 800
 QY 29 HisProCysAlaLeuAlaLeuIleGlyIleSerIleHisPro-----LeuCysPheAla 46
 DB 801 CACCCCAACATCATCATGAGCTGCGCGCGCGCTGTGCTGCAGAGCGGACCTCTGCTGTG 860
 QY 47 LeuGluLeuAlaProLeuSerSerLeuAsnThrValLeu---SerGluAsnAla----- 63
 DB 861 CTGAGATTGCGCGCGCGCGCGAGCGCTCAACCGAGCGCTGCGCGCGCGCGCGCGCG 920
 QY 64 -----ArgAspSerSerPheIleProLeuGluHisMetLeuThr 76
 DB 921 GACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT---CCGACGCTGTGTC 977
 QY 77 GlnLysIleAlaIleAlaIleAlaSerGlyLeuAlaIleIleHisLysLys----- 93
 DB 978 ---AACTGGCGCGCGCGAGTAGCGCGCGCGCATCTCTACTGTCATGAGCGCTTCGTG 1034
 QY 94 AsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValTrpSerLeuAspValLys 113
 DB 1035 CCCATCTCTCACCAGGAGCTCAAGTCCAGCAACATTTTCTGCTA-----CTTGAGAGATA 1088
 QY 114 GluHisIleAsnIle-----LysLeuSerSerPyrGlyIleSerArgGln 128
 DB 1089 GAGCATGATGACATCTGCAATAAACTTTGAGATTAACATTTTGGTTGGCGAGGAA 1148
 QY 129 SerPheHisGluGlnIleAlaLeuGlyValGluGlyThrProGlyTrpGlnAlaProGluIle 148
 DB 1149 TGCGACAGAGACCAAAATGAGCACAGCGGACCTATCCCTGATGCGCCCGAAGTG 1208
 QY 149 ArgProArgIleValIleTrpAspGluLysValAspMetPheSerTrpGlyMetValLeuTrp 168
 DB 1209 ATCAACTCTCTCTGTTTCTTAAGGAGAGGACATCTGAGCTATGAGAGTGTGTG 1268

QY 169 GluLeuLeuSerGlyGlnArgProAlaLeuGlnHisGlnLeuGlnIleAla----- 186
 DB 1269 GAAGTCTCATCCGAGAGAACTCCCTATCGGGGATGATGAGCGCTGCGCTGATGGG 1328
 QY 187 -----LysLysLeuSerLysGlyIleArgProValLeuGlnProGluGlnVal 203
 DB 1329 GTAGCATCAATTAATCTACTTTGCGCATTCATCCACCTGCGCTGACCGC----- 1379
 QY 204 GlnPheArgArgLeuGlnIleAlaLeuMetGluCysTrpAspThrLysProGluLysArg 223
 DB 1380 -----TTTGCAAGCTCATGAAAGATGTGGCAACAGACCTCATATTTCGT 1427
 QY 224 Pro---LeuAlaLeu-----SerValLysSerGln 232
 DB 1428 CCACTCTTTGCTTAATTTCTCCAGACTTGACTGCTATTGAGGGGACATGACTGAG 1487
 QY 233 MetLysAspProThrPheAlaThrPheMetLysGluLeuGlnCysGlyLysGlnThrAla 252
 DB 1488 ATGCCCTCAAGAACTTTTCATTCATCCAGAAAGATGACTGGAACTAGAAATTCACAAATG 1547
 QY 253 Phe-----PheSerSerGlnGlyGlnIleValTrp--- 262
 DB 1548 TTGATGAGTTGAGAAACAAAGAGAGCTGCGATCCCGGAGAGAGAGCTGACTCG 1607
 QY 262 ----- 262
 DB 1608 GCGGCTGTGCAGCAGAGTCTCAGAGAGAGCTGTAAGCGCGCTGAGCAGCTGCGCA 1667
 QY 263 -----ValValPheTrpAspGlyLys 269
 DB 1668 GAGCGGAGATGACAGCTGCTGAGCGGAGAACTTAACATTTCTGATTCACGATTAACCG 1727
 QY 270 GluGluSerArgAsnThrValValAsnThrGluLysGly----- 283
 DB 1728 GAGAAACCCAG-----GTAAAGAGAGAGAGGAGGCAAGTTTAAGAGAGTCTG 1775
 QY 284 ---LeuMetGluValAlaGlnArgMetCysCysProGly----- 294
 DB 1776 TTAACTCAACAAATGACATGACATGATTAATTAATTCATTCAGCAAGAAATTAAC 1835
 QY 295 MetLysValSerCysGlnLeuGlnValGlnArgSerIleThrAlaThrGluAspGln 314
 DB 1836 GTGCAAGCGCTCTCCAACTTGAGCAAAAGCGGAGCTGTAACAGCAGCACT----- 1886
 QY 315 LysIleTrpIleThrIleuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeu 334
 DB 1887 -----TCCAGTCCCGC----- 1898
 QY 335 AspThrProAlaValAlaThrCysPheLeuAlaValProValIleLysLysAsnSerTrp 354
 DB 1899 AGCAGCCCAACAAATGATGCGCGAGCTCCGAGCCATACAGTTCACATGTAAGC--- 1955
 QY 355 LeuValLeuAlaGlyLeuAlaAspGlyLeuValAlaValPhe----- 368
 DB 1956 -----AATAAACTTGGGAGAGGAAACAGCTTTTGGACAGAGAAATTTGAG 2003
 QY 369 ProValValArgGlyThrProLysAspSerCysSerTrpLeuGlnSerHisThrAlaAsn 388
 DB 2004 GATGTAAAGAAATTTTAAGAAAGATTTGATTCG-----GGA 2045
 QY 389 ArgSerLysPheSerIleAlaAspGlu---AspAlaArgGlnAsnProTrpProValLys 407
 DB 2046 CCNAATTCATTCAATTAAGATGAGACAGATTCGCAAGAAAGGATTAAGACT----- 2099
 QY 408 AlaMetGluValValAsnSerGlySerGluValTrpTrpSerAsnGlyProGlyLeuLeu 427
 DB 2100 -----CTCTCCGATGCGCAAGCTGTGTAACAT-----ACTTA 2135
 QY 428 Val-----IleAspCysAlaSerLeuGluIle-----Cys 437
 DB 2136 ATTAATAAATCAGAAACCAATGAGCTGCTCAATTTTGTGGACAGCAGGCTGCTGT 2195
 QY 438 ArgArgLeuGlnLupProTrpMetAlaProSerMetValThrSerValValCysSerSerGlu 457


```

Db 478 GACGGGGGGCGGCTCCGAGACGCGCGCGCGGCTCTGCGCATGCTGCGG 537
Oy 29 HisProCysIleValAlaLeuIleGlyIleSerIleHisPro-----LeuCysPheAla 46
Db 538 CACCCCAACATCATGAGCTGCGCGCGCGCTGCTGCACAGACCGGACCTGCGCGG 597
Oy 47 LeuGluLeuAlaProLeuSerSerLeuAnthrValLeu-----SerGluAsnAla----- 63
Db 598 CTGGAGTTCGCGCGCGCGCGAGCGAGCTGCTCAACGCGCTGCGCGCCACCGCGCG 657
Oy 64 -----ArgAspSerPheIleProLeuGlyHisMetLeuThr 76
Db 658 GACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
Oy 77 GlnIysIleAlaIleValIleAlaSerGlyLeuAlaIleValIleHisIleLys----- 93
Db 715 ---AACTGGCGCGCGCGAGATAGCGCGCGCATGCTCTACCTCAGAGAGAGCGCTCG 771
Oy 94 AsnIleIlePheCysAspLeuIleSerAspAnIleLeuValIleTrpSerLeuAspValLys 113
Db 772 CCCATCTCTGACCGCGAGCTCAAGCTCAAGCAACATTTCCTA-----CTTGAAGAATA 825
Oy 114 GluHisIleAsnIle-----LysLeuSerAspTyrGlyIleSerArgIn 128
Db 826 GAAATGATGATCATCTGCATATAAACTTTGAAGATTACGATTGGGTTGGCGAGGGA 885
Oy 129 SerPheHisGlyIleValLeuGlyValIleGlyIleThrProGlyIleAlaProGluIle 148
Db 886 TGGCAGACAGACCAAAATGAGCAGACGAGCAGCCTATGCTGATGCGCGCGCGAG 945
Oy 149 ArgProHisGlyIleValIleTyrAspIleLysValIleAspMetPheSerTyrGlyMetValLeuTyr 168
Db 946 ATCAAGTCTTCCTGTTCTTCTTAAGGAGAACGACATCTGAGCTATGAGATGCTGCTGG 1005
Oy 169 GluLeuLeuSerGlyIleValIleProAlaLeuGlyHisIleGlnLeuGlnIleAla----- 186
Db 1006 GAACAGCTCAGCGAGAACTCCCTATCGGCGCATGAGCGCTCGCGCGCTGATGG 1065
Oy 187 -----LysLysLeuSerLysGlyIleArgProValLeuGlyIleProGluIleVal 203
Db 1066 GTAGCAGTCAATAACTCACTTGGCCATTCATCCACCTGCGCGCGAGCG----- 1116
Oy 204 GlnPheArgIleGluGlnAlaLeuMetMetGlyCysTrpAspThrLysProGluLysArg 223
Db 1117 -----TTTGCACAGCTCATGAAGAATGCTGCAACAGACCTCATATTCTGT 1164
Oy 224 Pro-----LeuAlaLeu-----SerValIleSerGln 232
Db 1165 CCATGCTTTCCTTAATTCTCGAACAAGTGCATTCGTAAGGGGCGAGTGTGACTGAG 1224
Oy 233 MetLysAspProThrPheAlaThrPheMetTyrGluLeuCysGlyLysGlnThrAla 252
Db 1225 AGCCTCAAGATCTTTTCATTCATGCAAGATGACTGGAACCTGAATTCACAACAAG 1284
Oy 253 Phe-----PheSerSerGlnGlyGlnIleLysThr----- 262
Db 1285 TTGTAGTGAAGTGAACAAGAAAGAGAGCTGCGATCCGGAAGAGAGAGCTGACTCG 1344
Oy 262 ----- 262
Db 1345 GCGGCTCTCCACAGAAAGTCTCAGAGAGAGCTCTAAAGCGCGCTGAGACAGCTGSCA 1404
Oy 263 -----ValValPheThrAspGlyLys 269
Db 1405 GAGCGGAGATGAGCTGCTGAGCGGGAATTAACATCTGATATTCACAGTAACAAG 1464
Oy 270 GluGluSerArgAsnTyrThrValIleValAsnThrGluLysGly----- 283
Db 1465 GAGAAAGCCCAAG-----GTAAAGAAAGAGAAAGCAAGTTTAAGAAAGTCTGT 1512
Oy 284 -----LeuMetGluValGlnArgMetCysCysProGly----- 294

```

```

Db 1513 TTAAGCTCAAGATGAGACATGATTCAGTTTACTTCAGATTTCACACACAAGATAAC 1572
Oy 295 MetLysValSerCysGlnLeuGlnValGlnArgSerLeuThrPheAlaThrGluAspGln 314
Db 1573 GTGACGCGCTCTCCCAACTTGTGGACAAAGCGGAGCCTGAAACAGAGAGT----- 1623
Oy 315 LysIleTyrIleTyrThrLeuLysGlyMetCysProLeuAsnThrProGlnAlaLeu 334
Db 1624 -----TCCAGTCCCGCG----- 1635
Oy 335 AspThrProAlaValIleValIleThrCysPheLeuAlaValProValIleLysAsnSerTyr 354
Db 1636 AGCAGCCCGCACAAATATGATCCCGAGCTCCGAGCCATACAGTTACTTCAGATGAAGC--- 1692
Oy 355 LeuValIleAlaGlyLeuAlaAspGlyLeuValAlaValPhe----- 368
Db 1693 -----AATAAACTTGGGGAAGAGACACAGCTTTCGACAGAAGAAATTGAG 1740
Oy 369 ProValIleArgGlyIleThrProLysAspSerCysSerTyrIleCysSerHisThrAlaAsn 388
Db 1741 GATGTAAAGAAATTTAAGAAAGAGTTGTACTCG-----GGA 1782
Oy 389 ArgSerLysPheSerIleAlaAspGlu-----AspAlaArgGlnAsnProTyrProValLys 407
Db 1783 CCAATTCATTCATCAATGAAGATGAAACAGATTTCAGAAAGAGATAGACT----- 1836
Oy 408 AlaMetGluValIleValAsnSerGlySerGluValIleTyrIleSerAsnGlyProGlyLeu 427
Db 1837 -----CTCTCGATGCAACAGTCTGTGTCACAT-----ATCTTA 1872
Oy 428 Val-----IleAspCysAlaSerLeuGluIle-----Cys 437
Db 1873 ATAAAAAATCAGAAAAACCATGCGCTTGGCTTCATTTGTTGTGGACCGACCGAGTCTGT 1932
Oy 438 ArgArgLeuGluProTyrMetAlaProSerMetValThrSerValIleCysSerSerGlu 457
Db 1933 GAA-----GAGCCAAACTTCCCTGAT-----CGATTAGA 1965
Oy 458 GlyArgGlyGluGluValIleValIleValIleValIleValIleValIleValIleValIle 477
Db 1966 CACAGAAAAACCAAAATTAATTAATGCTGATGCTGACGCTTACATGATCTTGTGG 2025
Oy 477 HisSerThrThrTyrGlnLeuCysAlaArgTyrPheCysGlyValIleProSerProLeuArg 497
Db 2026 AAAGATGCTCAGAGAGAGAAATCTCAGAAAGCTGAAAGCTGGGAGAGAGACCTCTCG 2085
Oy 497 GAspMetPheProValIleArgProLeu-Asp-ThrGluProProAlaIleSerHisThrAla 516
Db 2086 A---ATGCTGCACAGTCTCTCATTCAGATGACTCTTACGAATAGCTGATGATCCCC 2142
Oy 517 AsnProLysValProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluGlu 536
Db 2143 CAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2202
Oy 537 LeuGlyThrGlnIle-----LeuIleHis----- 544
Db 2203 CTGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2262
Oy 545 -----GlnGluSerLeuThrAspTyrCysSerMetSerSerTyrSerSer 559
Db 2263 GAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2322
Oy 560 SerProProArgGln-----AlaAlaArgSerProSerSerLeuPro 573
Db 2323 AGTCTCCACAAAGCTGCTGCATCCACACCTGTGGAGAGAGAGAGAGAGAGAGAGAGAG 2382
Oy 574 SerSerProAlaSerSer-----SerSerValIleProPheSerHisAspCysGlu----- 589
Db 2383 CTGTCAAGTCCCTGGGATCCTCTCCACACCTCTTCTCCACAAAGAGAGAGAGAGAGAG 2442
Oy 590 -----AspSerAspMetLeu 594
Db 2443 ATGACAGTGAAGATCCACACTGTGTGACAGTGCACCTGTGACATTGTGAGATGCTC 2502

```

```

QY 522 uclyaspserllealaaspaValserlleMetYrSerGluGluLeuGlyThrGlnIleLe 542
DB 2496 CGGCGTGGGGCCCGGCGCC-----TCGGCCACCCCT 2525
QY 542 uillehtsGlnGluSerlleuThrAspTyrCysSerMetSerSerTyrSerSerProp 562
DB 2526 CCGTCGCGTGTCTCGCTCCGACGCAACCTCCAGCGCTCTCACTCTCGCTCTGACAG 2585
QY 562 ohrGlnAlaIala---ArgSerProSerSerleuProSerSerProAlaSerSerSer 581
DB 2586 TCACAGGGCCGACCGCGCCGCGCTCCACACCCCTCCGCGCGCCGACCCACCCAC 2645
QY 581 rValProPheSerThr-----AspCysGluAspSerAspMetLeuThrPr 597
DB 2646 GCCCTCGCCGACACCAACCCCTGTGTGACCTGTGAGCTGAGAGCTTCAAGAAGACCC 2705
QY 597 ogly-----AlaAlaSerAspArgSerGluH 606
DB 2706 CGGCGAGTGCCTACGCGCCACCCACGCTACGCGTGTATGCGTGTGAGCGCGG---CA 2762
QY 606 saspleuThrProMetAspGly 613
DB 2763 CCGCGCGACGCGCATCGATGG 2784

RESULT 7
US-09-938-842A-1014
Sequence 1014, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krieps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1014
LENGTH: 1662
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-1014

Alignment Scores:
Pred. No.: 2,01e-20 Length: 1662
Score: 270.50 Matches: 74
Percent Similarity: 48.36 Conservative: 44
Best Local Similarity: 30.338 Mismatches: 109
Query Match: 7.088 Indels: 17
DB: 9 Gaps: 8

US-09-836-392-21 (1-728) x US-09-938-842A-1014 (1-1662)
QY 16 GluPheArgGlnGluAlaSerMetLeuHsAlaLeuGlnHsProCysIleValAlaLeu 35
DB 946 GAATTTCTTCAGAGATATATATATAGAGAAAGTTCTGCGATATAAATGTTCACAGTTC 1005
QY 36 illeGlyIleSerIleHsPro-----LeuCysPheAlaLeuGlnLeuAlaProLeuSer 53
DB 1006 ATTGCGGATGTACAGATACCAACCTTGTGACAGAGTTGACATGATCGGGGG 1065
QY 54 SerLeuAsnThrValLeuSerGluAsnAlaArgAspSerSerPheIleProLeuGlyHs 73
DB 1066 AGCATTTATGATTCCTTCACAAACAC-----AAAGGGGTTTAAATTCATCATCT 1116

```

```

QY 74 MetLeuThrGlnIleValAlaTyrGlnIleAlaSerGlyLeuAlaTyrLeuHsIleLys 93
DB 1117 TTGCTC-----AAAGTGCACCTCGACGCTCTGCAAGAGATATATCTGCATCAAAAC 1170
QY 94 AsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValTrpSerLeuAspValLys 113
DB 1171 AATATATATTCANAGACACCTTAAGACTGCTAATCTTCT-----ATGAGAC 1215
QY 114 GlnHsIleAsnIleLysLeuSerAspTyrGlyIleSerArgGlnSerPheHsIleGly 133
DB 1216 GAACATGAAGTGTCTCAAGTGTCCGATTTGGTGTGCCAGAGTGCAGACTGATGAGG 1275
QY 134 AlaLeuGlyValGlu---GlyThrProGlyTyrGlnAlaProGluIleArgProArgIle 152
DB 1276 GTTATACAGCGGAAGAACAGGACATCCAGTGTATGCTCCAGAGTCAATTGACACAAA 1335
QY 153 ValTyrAspGluLysValAspMetPheSerTyrGlyMetValLeuTyrGluLeuLeuSer 172
DB 1336 CCTTATGATACACAGGCGCAGATGTCTTACACTACGCGCATGTGCTGTGGAACTTTGACT 1395
QY 173 GlyIleArgProAlaLeuGlnHsIleGlnIleAla---LysLysLeuSerLys 191
DB 1396 GGGGAACCTCCCATATTTCTTACTGACTCCACTGCAAGCTGTGTGGCGTTGCCAAAG 1455
QY 192 GlyIleArgProValLeuGlnIleProGluGluValGlnPheArgTrpLeuGlnAlaLeu 211
DB 1456 GGACTTAGACCAAAAAT-----CCAAAGGAACACAC---CCAAACTGACTGACTT 1506
QY 212 MetMetGluCysTrpPheThrLysProGluLysArgProLeuAlaLeuSerValSer 231
DB 1507 CTGAGAAATAGCTGCGACAGAACCCAGCTCTAAGACCCATTTGCAGAAATCATAGAA 1566
QY 232 GlnMetLysAspProThrPheAlaIleThrPheMetLysGluLeuCysGlyLysGlnThr 251
DB 1567 ATGCTTACCACTAACTCCGAGGTAAATTGATTATCATTTGATTAAGATAACATGCT 1626
QY 252 AlaPhePheSer 255
DB 1627 GGTACTTTTCA 1638

RESULT 8
US-10-014-882-1
Sequence 1, Application US/10014882
Patent No. US20020107384A1
GENERAL INFORMATION:
APPLICANT: Hu, Yi
APPLICANT: Kleke, James
APPLICANT: Donoho, Gregory
TITLE OF INVENTION: No. US20020107384A1el Human Kinase and Polynucleotides Encod
FILE REFERENCE: LEX-0279-USA
CURRENT APPLICATION NUMBER: US/10/014,882
CURRENT FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: US 60/254,744
PRIOR FILING DATE: 2000-12-11
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 3111
TYPE: DNA
ORGANISM: homo sapiens
US-10-014-882-1

Alignment Scores:
Pred. No.: 7.1e-20 Length: 3111
Score: 269.50 Matches: 177
Percent Similarity: 36.66 Conservative: 102
Best Local Similarity: 23.266 Mismatches: 258
Query Match: 7.058 Indels: 224
DB: 12 Gaps: 36

US-09-836-392-21 (1-728) x US-10-014-882-1 (1-3111)
QY 9 AspAlaMetLysAsnPheserGluPheArgGlnGluAlaSerMetLeuHsAlaLeuGln 28

```

NUMBER OF SEQ ID NOS: 318
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 226
 LENGTH: 3454
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-969-347-226

Alignment Scores:

Pred. No.:	4,35e-24	Length:	3454
Score:	307.50	Matches:	178
Percent Similarity:	35.43%	Conservative:	87
Best Local Similarity:	23.80%	Mismatches:	272
Query Match:	8.05%	Indels:	212
DB:	10	Gaps:	31

US-09-836-392-21 (1-728) x US-09-969-347-226 (1-3454)

```

QY 19 GNGIUALASerMetLeuNHISAlaLeuGlnHISProCysIIleValAlaLeuIIleGlyIle 38
DB 718 CAGGAAGCCCGGCTCTTGGAGCCCTGCAGACACCCACATATATCCCTTAGGGCCG 777
QY 39 SerIIeHISPro-----LeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsn 56
DB 778 TGCCTCAACCCCGCCACACCTCTGCTAGTAGAGTAGATGAGCGGGGTGGTGCACAGAC 837
QY 57 ThrValIleSerGlnuSnaAlaArgAspSerPheIIleProLeuGlnHISMetLeuThr 76
DB 838 AGGGTCTGGCAGGTGCGCG-----GTGCA-----CCTCAGCTGTGTC 879
QY 77 GlnLysIIleAlaTyrGlnIIleAlaSerGlyLeuAlaTyrLeuHISLysLys----- 93
DB 880 ---AACTGGGCTGTGACAGGTGGCGGCGGACATGACTACTACACATGATGCCCTGTG 936
QY 94 AsnIIleIlePheCysAspLeuLysSerAspAsnIIleLeuValTyrSerLeuAspValLys 113
DB 937 CCCATATCCACCGGAGCCTCAATCATCATCATCTGATC-----CTGGAGGCCATC 990
QY 114 GlnHISIIleAsn-----IIleLysLeuSerAspTyrGlyIleSerArgGln 128
DB 991 GAGAACACCAACTGTCGACACAGCGGTGCTCAAGATCAGCGACTTCGGCTCGCCCGCCAG 1050
QY 129 SerPheHISGlnLysAlaLeuGlnLysAlaGlnGlyThrProGlyTyrGlnAlaProGlnIle 148
DB 1051 TGGCACACAGACACCAAGATGAGCGCTGCGGGACCTACGCGTGATGGCGCGGAGGT 1110
QY 149 ArgProArgIIleValTyrAspGlnLysValAspMetPheSerTyrGlyMetValLeuTyr 168
DB 1111 ATCCGCTCTCCCTCTCTCCCAAGACAGTATGATCTGAGAGCTTCGGGTGCTGTGG 1170
QY 169 GlnLeuLeuSerGlyLysIleArgProAlaLeuGlnHISGlnLeuGlnIIleAla----- 186
DB 1171 GAGCTCTCAGCGGGAGAGTCCCTTACCGTGAATGAGAGCCCTTGGCGCTGGCTATGGC 1230
QY 187 -----LysLysLeuSerLysGlyIleArgProValLeuGlnLysGlnProGlnLysAla 203
DB 1231 GTGGCTATGAATTAAGCTGAGCTGCTCCATTCCTCCACAGCTGCCGCGAGCC----- 1281
QY 204 GlnPheArgArgLeuGlnAlaLeuMetGlnCysTyrAspThrLysProGlnLysArg 223
DB 1282 ---TTTGGCCGCG-----CTCCTGAGAGAACTGTGGAGCCACAGCCCGCGGCG 1329
QY 224 ProLeuAlaLeuSerValAlaSerGlnMetLysAspProThrPheAlaThrPheMetTyr 243
DB 1330 -----CCAGATTTCGGTACATCTTGAG 1353
QY 244 GlnLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnIleTyrThrVal 263
DB 1354 CGGCTTGAAGTATCGAAGAGTACAGCTTTCAGATGCCATCGAGAGCTTCACCTCG 1413
QY 264 ValPheThrAspGlyLysGlnLysArgAspTyr-----ThyValValAsnThrGln 281
DB 1414 CTGCAGGAAGACTGGAAGCTGGAGATTCAAGACATGTTTGTGATGACCTTGCAGCAAGAG 1473

```

```

QY 282 LysGlyLeuMetGln-ValGlnArgMetCysCys-----ProGlyMet 295
DB 1474 AAGGAGCTTCGAGAGCCGTGAAGAGAGTGTGCGGGCGGACAGAGAGCGCTTCAG 1533
QY 295 LysValSerCys-----GlnLeuGlnValGlnArgSerLeuTyrThrAlaThrGln 312
DB 1534 GAGGAGCAGCTGCGGGCGGGGAGCAGAGCTGGCGAAGCGTAGATGACAC----- 1585
QY 312 uAspGlnLysIleTyrThrLeuLysGlnMetCysProLeuAsnThrProGlnIle 332
DB 1586 -----TCGTGGAACGGAGACTGCACCTGCTCATGTGCAGCTG 1623
QY 332 nAlaLeuAspThrPro----- 337
DB 1624 AGCCAGAGAGAGCCCGGGTCCGCAAGCGCAAGCCAACTTCAGGCGCGCTGCTC 1683
QY 338 -----AlaValValThrCysPheLeuAlaValProValIIleLysLys 351
DB 1684 AAGCTGCGGGAAGCGGCGGACCATCATCAGCTGCTGCTTGGATGATGATGAC 1743
QY 351 sAsnSerTyrLeuValLeuAlaGlyLeuAlaAspGlyLeuValAlaValPhePro----- 369
DB 1744 GTCCAGGCTCTCCAACTCTG-GATTAAGCGGAAGGATCCGATGGGCGCGCCCTGC 1802
QY 370 -----ValValArgGlyThrProLysAspSerCysSe 380
DB 1803 AAGCCCAACATCATCCCGGCTGAGGCGCATTCCTGACTCCGTGAC----- 1854
QY 380 rTyrLeuCysSerHISThrAlaAsnArgSerLysPheSerIIleAlaAspGluAspAlaArg 400
DB 1855 -----TGTGTGGAGCAGCAGTGGCAGCAGCATGTGAGAGAGTGGACATGAGCGCG 1907
QY 400 gGlnAsnProTyrProValLys----- 407
DB 1908 CGGTGGGCGCC---CCAAAGAGGAAGAACTGTGCGGGGCAAGAAAGAGAGCAACGTG 1964
QY 408 -----AlaMetGlnValVal 412
DB 1965 GGGGCCAGCTCCACCCCTCAGAGAGAGCGGTGGAGAGAGAGAGCTGAAGGGCT 2024
QY 412 IAsnSerGlySerGlnValTyrTyrSerAsnGlyProGlyLeu----- 426
DB 2025 GGGGAGAGGAACAAACAGTGTGATCAAGTGCACCCCAACGTGGGCAATCCCCAAACA 2084
QY 427 -----LeuValIIleAspCysAlaSerLeuGlnIleCysArgArgLeu----- 440
DB 2085 CACACCCATCGCCCTGCTGCGCCAGCCTCAATGAGATGAGAGATTGCCGAGGACAGA 2144
QY 441 -----GlnProTyrMetAlaProSerMetValThrSerVal 452
DB 2145 GGATGAGGACAGACAGCGTCCCTTCCCTTACCTGAGACCCGCTTACCTTCACTAGTGC 2204
QY 452 IValCysSerSerGlnGlyArgGlyGlnIleValValTyrCysLeuAspAspLysAlaAs 472
DB 2205 ACTGCTCGGAGACCTCCCGCGGGGCGGGCGCGGTGGAGCGGACGCGCTCGCGCG 2264
QY 472 nSerLeuValMetTyrHISerThrThrTyrGlnLeu-----CysAl 486
DB 2265 CCCCCTGCGTGGGAGACAGCGCGCGCGGCGGTGCGACCTGCGGCGCTGAGCTGGCG 2324
QY 486 a----- 486
DB 2325 CACGCTGCTGGGGGCTGTGGGCTGCGGCCACAGCTGCGCGAGCGCGGCGCGAGCG 2384
QY 487 -----ArgTyrPheCysGlyValProSerProLeuArgAspMetPheProVal 502
DB 2385 TGAGAGACAGCGCGCTGCTGAGCGCTTCTTCTTCCCGCGCGCGCTTCGCG- 2442
QY 502 IArgProLeuAspThrGlnProProAlaAlaSerHISThrAlaAsnProLysValProGln 522
DB 2443 -CGGGGCGCTC-----AGCCACACCGCGCGTCCACAGCGCGCGCGAGAGAGAGTGGCGCC 2495

```

Db 4656 TACCAGATCCGCTCGGCGCTGACCTACCTGCACAGAAAACATCATCTTGTGACCTG 4715
 Qy 101 LysSerAspAsnIleuValItrpSerIeuAspValIysGluHisIleAsnIleLysLeu 120
 Db 4716 AAGTCGACAAACATTCGTGGTGTGCTCCCTGAGCTCAAGAGGACCATCAATCAACTA 4775
 Qy 121 SerAspTyrGlyIleSerArgIleSerPheHisGluGlyAlaLeuGlyValGlyGlyThr 140
 Db 4776 TCTGACTACGCGGATTCAGAGCATTCATTCATGAGCGCGCTGAGGCGGCTGAGGCGACT 4835
 Qy 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet 160
 Db 4836 CTTGGCTACCAAGCGCCAGAGATAGGCTCTGCATTTATATGATGAGAGATGATATG 4895
 Qy 161 PheSerTyrGlyMetValIleuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlyHis 180
 Db 4896 TTCTCCATGGAAATGGTCTGCTACAGATTGCTGCAGAGACGCGCTGACCTGGCGCAC 4955
 Qy 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGlnPro 200
 Db 4956 CACAGCTCCAGATTGCCAAGAACCTCTCCAAAGGCGATCCGCGGTCTGCGGACAGCG 5015
 Qy 201 GluGlnValGlnPheArgArgLeuGlnAlaLeuMetGluCysTyrPaspThrLysPro 220
 Db 5016 GAGAAATGCAATTCGCGCGCTCAGAGCGCTCATGATGAGTCTGGAGACTGAGCA 5075
 Qy 221 GluLysArgProLeuAlaLeuSerValIleSerIleMetLysAspProThrPheAlaThr 240
 Db 5076 GAGAAAGGACCGCGCGCTGCTGCTGAGGAGCAGATGAGAGACCGCACTTTGGCCACC 5135
 Qy 241 PheMetTyrGluLeuCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGlu 260
 Db 5136 TTCATGTATGAACTGTCTGTGGAGACAGACAGCTTCTCATCCAGCGGACAGAG 5195
 Qy 261 TyrThrValIlePheTyrPaspGlyLysGlnLeuSerArgAsnTyrThrValIleAsnThr 280
 Db 5196 TACACCGGTGTGTGGAGAGAAAGAGAGTCCAGAACTACACGCTGTGTAACA 5255
 Qy 281 GluLysGlyLeuMetGluValIleArgMetCysCysProGlyMetLysValSerCysGln 300
 Db 5256 GAGAAAGGCGCTCATGAGAGTGCAGAGATGTGCTGCGCTGGAGTGAAGTACAGTCCGAC 5315
 Qy 301 LeuGlnValGlnArgSerLeuTyrThrAlaThrGluAspGlnLysIleTyrIleTyrThr 320
 Db 5316 CTCAGAGTCCAGAGATCCCTGTGGACAGCCACCGAG 5351
 Qy 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
 Db 5351 ----- 5351
 Qy 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValIleuAlaGlyLeu 360
 Db 5352 -----AATTCCTACCTGTGTAGCGGCGCTC 5378
 Qy 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 Db 5379 GCCGATGGCGCTGTGGCTGTTCCTCCGTGTGGCGGACCCCAAGAGACAGCTGTCC 5438
 Qy 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
 Db 5439 TACCTGTGCTCACACACAGCCACAGGTCACAGTCCAGATCCGCGAGAGAACCCACAG 5498
 Qy 401 GlnAsnProTyrProValLysAlaMetGluValValAsnSerGlySerGluValIleTyrPyr 420
 Db 5499 CAGAAACCCCTACCCAGTGAAGCCATGAGAGTGTCAACAGCGGCTTCTGAGGTCTGCTGAC 5558
 Qy 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
 Db 5559 AGCAATGGCGCGGCTCTCTGTATGACATGTGCTCCCTGGAGAAATGTGACAGGCGGCTG 5618
 Qy 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGluLysArgGly 460

Db 5619 GAGCCCTACATGGCCCCCTCATGTTAGCTCAGTGTGTGACAGTCTGAGGACAGAGG 5678
 Qy 461 GluGlnValValIleTyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 Db 5679 GAGAGAGTGTCTGTGCTCTGATGATGACAAAGCCAACTCTGTGTGATGTACCACTCCACC 5738
 Qy 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuAspMetPhe 500
 Db 5739 ACCTACAGAGTGTGTGCGGCTGCTTCTGTGGGGGTCCCAAGCCCCCTCAGAGGACATGTTT 5798
 Qy 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
 Db 5799 CCGGTGGCGCTTGTGACAGAGAACCCCGGACCCAGCCACAGCGGCAACCCAAAGTGTCC 5858
 Qy 521 ProGluLysAspSerIleAlaAspValSerIleMetTyrSerGluLeuGlyThrGln 540
 Db 5859 CCTGAGGGGAGCTCATCGGAGAGTGTGAGCATGTATGATGAGTGTGAGGAGCGGACAGCAG 5918
 Qy 541 IleLeuIleHisGlnGlnLeuSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 560
 Db 5919 ATCTGTATCCACCGAGATCATCTGACTACTGCTCATGTCTCTCTACTCTCATGCC 5978
 Qy 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
 Db 5979 CCACCCCGGACAGCTGCGCAGAGTCCCTCAAGCTCCCAAGCTCCCAAGCAAGTCTTCC 6038
 Qy 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
 Db 6039 AGTGTGCTTCTTCCACCGAGTGTGAGAGACATGACATGTATGATGAGCGGCTGTGCC 6098
 Qy 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeu 620
 Db 6099 TCCACAGAGTGTGAGCATGACCTGACCCCAATGAGCGGAGAGACTTTCAGCGACACCTG 6158
 Qy 621 GlnAlaValLysIleLeuAlaValAlaArgAspLeuIleTyrValProArgArgGlyLysP 640
 Db 6159 CAGGCGGTGAAGATCTCTCGCGGTGAGAGACTCTTTGGGTGCCAGCGCGGTGAGAT 6218
 Qy 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 Db 6219 GTTATGTGATGTGCTGTGAGAAAGATTTCTGCGCCAGCGGCGCCAGTCAATGCGCTC 6278
 Qy 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 680
 Db 6279 TTAAGACCCCGAGAGCTGACCTCCAGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6338
 Qy 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrCysLeuAlaValIleArg 700
 Db 6339 GACACTGTGTGTGACACTTTGAAATGAATAACACAGATGTGTGTGTGTGTGTGTGTGT 6398
 Qy 701 GlyTyrGlyAlaArgGluPheAspIlePheTyrGlnSerTyrGluGluLeuGlyArgLeu 720
 Db 6399 GCGTGGGGCGGACGAGAGTGTGACATTTTACACAGTCTTACAGAGAGTGTGGCGGCTG 6458
 Qy 721 GluAlaCysThrArgLysArgArg 728
 Db 6459 GAGGCTTGCACTCCCAAGAGAG 6482

RESULT 6
 US-09-969-347-226
 Sequence 226, Application US/09969347
 Patent NO. US20020115085A1
 GENERAL INFORMATION:
 APPLICANT: Ebner, Reinhard
 TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using S
 TITLE OF INVENTION: Sets
 FILE REFERENCE: 689290-69
 CURRENT APPLICATION NUMBER: US/09/969,347
 CURRENT FILING DATE: 2001-10-02
 PRIOR APPLICATION NUMBER: US/60/237,598
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: US/60/237,604
 PRIOR FILING DATE: 2000-10-03

```

QY 241 pheMetTyrGluLeuCyCysGlyLysGlnThrAlaPhePheSerSerGlnGlnGlu 260
DB 5058 TTCATGTATGAAGTGTGCTGGGAAGCAGACAGCTTCTTCATCCAGGAGCCAGAG 5117
QY 261 TTTTThValValPheThrPaspGlyLysGlnGluSerArgSerTyrThrValValAsnThr 280
DB 5118 TACACCGTGTGTTTGGATGGAAAGAGAGTCCAGACTACAGGCTGTGTGAACACA 5177
QY 281 GAluYsgLYLeuMetGluValGlnArgMetCysCysProGlnMetLysValSerCysGln 300
DB 5178 GGAAGAGGCTATGAGAGGTGAGAGAGATGCTGCTGGATGAGAGGTGAGCTGCCAG 5237
QY 301 LeuGlnValGlnArgSerLeuTyrThrAlaThrGluAspGlnLysIleTyrIleTyrThr 320
DB 5238 CTCACAGGTCCAGAGATCCCTGTGGACAGCCAGAG----- 5273
QY 321 LeuLYsgLYMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
DB 5273 ----- 5273
QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
DB 5274 -----AATTCCTACTGTGCTTACCGGGCTTC 5300
QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyLysProLysAspSerCysSer 380
DB 5301 GCCGATGGGCTGTGCTGTGTTCCCGTGTGGGGGAGACCCCAAGAGACAGCTGCTCC 5360
QY 381 TyrLeuCySerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
DB 5361 TACCTGTGTACACACACAGCCCAAGTCCAGATCGCGGATGAGACGCGACGG 5420
QY 401 GlnAspProTyrProValLysAlaMetGluValValAsnSerGlySerGluValTyrPyr 420
DB 5421 CAGAACCTTACCACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5480
QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnLysArgArgLys 440
DB 5481 AGCAATGGGCGGGCTCTGTGTCACAGACTGTGCTCCCTGAGAGATCGACAGCGGCTG 5540
QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGluLysArgGly 460
DB 5541 GAGCCCTACATGGCCCTTCATAGTGTACGTAGTGTGAGCTGAGGCGCAGAGGG 5600
QY 461 GluGluValValTyrCysLeuAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
DB 5601 GAGGAGGTGTGTGTGCTGTGATGACAAAGCCAACTCTGTGTGTATACCACTCCACC 5660
QY 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
DB 5661 ACCTACACACTGTGTGCGGAGTACTCTGCGGGGTCCCAAGCCCTCAGGAGCAATGTT 5720
QY 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
DB 5721 CCCGTGCGGCTGTGAGACAGGAAACCCCGGAGCCAGCCACAGCCCAAAAGGAGTG 5780
QY 521 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGluGlnGlyThrGln 540
DB 5781 CCGTGAAGGAGACTCCATCGCGGAGTGAACATCATATACAGTGAAGAGTGGGCGAGGAG 5840
QY 541 IleLeuIleHisGlnGluSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSer 560
DB 5841 ATCTGATCCACAGAGAAATCACTGACTGCTCATGTCCTCATGCTCCATCTCATCC 5900
QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
DB 5901 CCACCCCGGAGGCTGCGAGGTCCCTCAAGCTCCCAAGCTCCCAAGCTCCCAAGTGTTC 5960
QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
DB 5961 AGTGTCCTTTTCCACCGACTGCGAGACTCAGACATGCTACATAGCCCGGCTGCTGCC 6020

```

```

QY 601 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
DB 6021 TCCGACAGGTCTGACAGCAATGACTGACCCCATGAGCGGGAGAACCTTCAGCCAGCACTG 6080
QY 621 GlnAlaValLysIleLeuAlaValArgAspLeuIleTyrPvalProArgArgGlyLysAsp 640
DB 6081 CAGGCGGTGAAGATCTCTGCGCTGACAGACCTCATATTGGGTCCCAAGCGCGCTGAGACT 6140
QY 641 ValIleValIleGlyLeuGlnLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
DB 6141 GTTATCGCATTTGGCTGTGAGAGAGATTTGCGCGCCAGCGGGCGCGAGCTATTCGCTC 6200
QY 661 LeuLYsAlaArgGluLeuThrProHisGlyValLeuValAspAlaValAlaLys 680
DB 6201 TTTAAAGCCCGGAGACTGACTCCGCATGGGGGTGTGTGATGTCTGCGGTGGCAAG 6260
QY 681 AspThrValValCysThrPheGlnAsnGlnAsnThrGluTyrCysLeuAlaValTyrArg 700
DB 6261 GACACTGTGTGTGTCACCTTGTCAAAATGAAACACAGAGTGTGCTGCTGCGCTGAGG 6320
QY 701 GlyTyrPglValAlaArgGluPheAspIlePheTyrGlnSerTyrGluGlnLysArgLeu 720
DB 6321 GGTGGGGGCGCAGGAGGTGCAATTTTCTACAGTCTTACGAGAGCTGGGCGGCTG 6380
QY 721 GlnAlaCysThrArgLysArgArg 728
DB 6381 GAGGCTTGCACTCGCAAGAGAGG 6404

RESULT 5
US-10-132-382-3
; Sequence 3, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 7307
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-382-3

Alignment Scores:
Pred. No.: 0 Length: 7307
Score: 3579.50 Matches: 688
Percent Similarity: 94.51% Conservative: 0
Best Local Similarity: 94.51% Mismatches: 1
Query Match: 93.66% Indels: 39
DB: 9 Gaps: 1

US-09-836-392-21 (1-728) x US-10-132-382-3 (1-7307)
QY 1 MetLeuArgHisLeuArgLarThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 20
DB 4416 ATGCTGAGGACCTCGGGGCGCCAGATGCGCATGAAGAACTTCTCCAGTCCGGCAGAG 4475
QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyLysSerIle 40
DB 4476 GCCAGAGTCTGTCAGCGGTGACAGACCCCTGCATCGTGGCGCTCATCGCATGAGCAG 4535
QY 41 HisProLeuCyPheAlaLeuGlnLysLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
DB 4536 CACCGGCTGTGCTGCGCTGAGGCTCGCGCGCTCAGCAGCAGCTCAACCGCTGTGTC 4595
QY 61 GlnAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 80
DB 4596 GAGAAGCCAGAGATTTCTTATACCCCTGGAGACATGCTACCCAAAAATAGCC 4655
QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100

```

```

Db 5225 TACCTGTGTCACACACAGCAAGAGTCCAACTTACAGATCGCGATGAGAGCGACG 5284
Oy 401 GlnAsnProTyrProValIysAlaMetGluValIAsnSerGlySerGluValITPyr 420
Db 5285 CAGAAACCCCTACCCAGTGAAGGAGCATGAGTGTGTCAACAGCGCTGTGAGTGTAC 5344
Oy 421 SerAsnGlyProGlyLeuValIleAspCysAlaSerLeuIleGlySerGlyLeu 440
Db 5345 AGCAATGGGCGGCGCTCTGTGATCAGTGTGCTCTCCCTGAGATCGACAGCGGCTG 5404
Oy 441 GluProTyrMetAlaProSerMetValIThrSerValIAsnSerGlySerGlyLeu 460
Db 5405 GAGCCCTACATGCCCCCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5464
Oy 461 GluGluValValITrpCysLeuAspAspGlyAlaAsnSerLeuValIleTyrHisSerThr 480
Db 5465 GAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5524
Oy 481 ThrTyrGluLeuGlyAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
Db 5525 ACCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5584
Oy 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProGlyVal 520
Db 5585 CCGGTGCGGCGCTTGGACAGAGAACCCCGGACGACGACGACGACGACGACGACGACG 5644
Oy 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluGluGlyThrGln 540
Db 5645 CCTGAGGGGAGCTCCATCGGAGCTGAGCATCATGATGATGATGATGATGATGATGATG 5704
Oy 541 IleLeuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 560
Db 5705 ATCTGATCCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5764
Oy 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
Db 5765 CCACCCCGGACGCTGCGAGAGTCCCGCTCAAGCTCCCGACGCTCCCGACGAGTGTTC 5824
Oy 581 SerValProPheSerThrAspCysGlyAspSerAspMetLeuHisThrProGlyAlaAla 600
Db 5825 AGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5884
Oy 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
Db 5885 TCCGACAGGTCTGAGCATGACCTGACCCCATGACGCGGAGACCTTCACGCGACACTG 5944
Oy 621 GlnAlaValIysIleLeuAlaValArgAspLeuIleTyrValProArgArgGlyIAsp 640
Db 5945 CAGGCGGTGAAGATCTCGCGCTCAGAGACCTCATTTGGGTCCCGGCGGCGGTGAGAT 6004
Oy 641 ValIleValIleGlyLeuGluIlyAspSerGluAlaGlnArgGlyValIleAlaVal 660
Db 6005 GTTATCTGTATTTGGCTCGGAGAGAGATTTGGCGCCAGCGGCGCCAGTCTATTTGCG 6064
Oy 661 LeuIysAlaArgGluLeuThrProHisGlyValIleValAspAlaAlaValIleAlaIlys 680
Db 6065 TTAAAGCCCGAGAGCTGATCCGATGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6124
Oy 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrCysLeuAlaValITrPArg 700
Db 6125 GACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6184
Oy 701 GlyTyrPheIlyAlaArgGluPheAspIlePheTyrGlnSerTyrGluGluLeuGlyArgLeu 720
Db 6185 GGGCGGGGCGCGAGGAGTGTGACATTTTCTACAGTCTTACAGAGAGCTGGGCGGCTG 6244
Oy 721 GluIlyAspThrArgIlyAspArg 728
Db 6245 GAGGCTTGCACTCGCAAGAGAGG 6268

```

```

; Sequence 7, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 7229
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-132-382-7

Alignment Scores:
Pred. No.: 0 Length: 7229
Score: 3579.50 Matches: 688
Percent Similarity: 94.51% Conservative: 0
Best Local Similarity: 94.51% Mismatches: 1
Query Match: 93.66% Indels: 39
DB: Gaps: 1

US-09-836-392-21 (1-728) x US-10-132-382-7 (1-7229)

Oy 1 MetLeuArgHisIleuValArgAlaThrAspAlaMetIlyAspPheSerGluPheArgGlnIu 20
Db 4338 ATCTAGAGCACTGTGGGCGGACCGAGTCCATGAGAACTTCCGAGTTCGCGAGAG 4397
Oy 21 AlAspMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
Db 4398 GCCAGATGCTGACCGCGGCGGACGACACCCCTGATGTGGCGCTCATCGGCATCAGCAG 4457
Oy 41 HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerIleuAsnThrValIleuSer 60
Db 4458 CACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4517
Oy 61 GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnIlyAla 80
Db 4518 GAGAGCCCGAGAGATTTCTTTATACCTTGAGACATGTGTCAACCCCAAAATATGCC 4577
Oy 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisIlyAsnIleIlePheCysAspLeu 100
Db 4578 TACCAGATGCGCTGTGGCGCTGACCTGACCTGACACAGAAATCATCTGTGTGACTG 4637
Oy 101 LysSerAspAsnIleuValITrPheSerLeuAspValIlySerGlnHisIleuIlyLeu 120
Db 4638 AAGTCGGACAAATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4697
Oy 121 SerAspTyrGlyIleSerArgIleSerPheHisGlyGluIlyAlaLeuGlyValIleGly 140
Db 4698 TCTGACTACGGGATTTTGAAGCATCTTTCATGAGGCGCCCTAGGCGTGGAGGGGACT 4757
Oy 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValITyrAspGlyValIAspMet 160
Db 4758 CCGGCTACAGAGCGCCAGAGATCAGGCTCGCATTTATATATGATGAGAGTATATG 4817
Oy 161 PheSerTyrGlyMetValIleuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlyHis 180
Db 4818 TTCTCTTATGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4877
Oy 181 HisGlnLeuGlnIleAlaIlyLysLeuSerLysGlyIleArgProValIleGlyGlnPro 200
Db 4878 CACAGCTCCAGATTTGCCAAGAACTTCCAAAGGCGATCCGCGCGGTCTTGCGGAGCG 4937
Oy 201 GluIlyValGlnPheArgArgLeuGlnAlaLeuMetMetIlyCysTrpAspThrIlyPro 220
Db 4938 GAGGAATGCAATTCGCGGCGAGCTCAGGCGCTATATGAGTGTGTGTGTGTGTGTGTGT 4997
Oy 221 GluIlyAspProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThr 240
Db 4998 GAGAAAGCAGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5057

```



```

|||||
Db 2062 GACACTGTGTGTGACCTTGAATAAAGAAACAGAGTGCTGCGCGCTGGAGG 2121
Oy 701 GYTTPGLAAlaargluPheaspIlePheTyrGlnSerTyrGluGluLeuGlyArgLeu 720
Db 2122 GGCTGGGGCGCCAGAGGAGTTTGCACATTTTTCACAGTCTTCAGAGAGAGTGGCGGCTG 2181
Oy 721 GluAlaCysThrArgLysArgArg 728
Db 2182 GAGGCTTGCACTCGCAAGAGAGG 2205

RESULT 2
US-10-132-382-5
; Sequence 5, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 7015
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-382-5

Alignment Scores:
Pred. No.: 0 Length: 7015
Score: 3579.50 Matches: 688
Percent Similarity: 94.51% ConservatIve: 0
Best Local Similarity: 94.51% Mismatches: 1
Query Match: 93.66% Indels: 39
Gaps: 1

US-09-836-392-21 (1-728) x US-10-132-382-5 (1-7015)
Oy 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 20
Db 4124 ATGCTGAGGACACTGCGGGCCACCGATGCATGAAGAACTTCTCGAGTCCGGCAGAGG 4183
Oy 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
Db 4184 GCCAGCAGCTGCACGCGCTGCAGACCCCTGCATCGCGCTATGGCATCAGCAGATC 4243
Oy 41 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
Db 4244 CACCGGCTCTGCTGCGCTGCAGCTCGCGCGCTCAGACAGCTCAGACCGTGTCTCC 4303
Oy 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 80
Db 4304 GAGAAAGCCAGATGTTCTTCTTATACCCCTGCGACACATCTCACCAAAAAATAGCC 4363
Oy 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
Db 4364 TACCAAGATGCGCTCGCGCTGCGCTTACCTGCACAAAGAAATACATCTTCTGTGACCTG 4423
Oy 101 LysSerAspAsnIleLeuValIlePheSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
Db 4424 AAGTGGAGAACATCTCTGTGTGTGCTTGTGACGTCAAGAGACACATCAACATCAAGCTA 4483
Oy 121 SerAspTyrGlyIleSerArgGlnSerPheHisGluGlyAlaLeuGlyAlaGluGlyThr 140
Db 4484 TCTGACTACGGGATTTCCAGCGAGATTCATTCATGAGGGCGCCCTAGCGCTGAGAGGCACT 4543
Oy 141 ProGlyTyrGlnAlaProGluIleArgProArgGlyIleValTyrAspGluLysValAspMet 160
Db 4544 CCGGCTTCCAGGCCCCAGATCAGGCTCGCATTTGTAATGATGAGAGAGGTAGATATAG 4603
Oy 161 PheSerTyrGlyMetValLeuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlnHis 180
|||||

Db 4604 TTCTCTATGGAATAGTGCTCTACAGAGTGTCTGACAGACAGCCCTGCATCGGCCAC 4663
Oy 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGlnPro 200
Db 4664 CACCAAGCTCCAGATTCGCAAGAGACTGTCCAAAGGACATCCGCCGGTTCGTGGGACACCG 4723
Oy 201 GluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGlyCysTrpAspThrLysPro 220
Db 4724 GAGAAAGTGCAGTTCGGCGGACGCTCAGCGGTATGATGAGTCTGGAGACTAAGACCA 4783
Oy 221 GluLysArgProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThr 240
Db 4784 GAGAAAGCAGACCGCTGCGCTGCTCGGTGTGAGCCAGATGAAGAGCCCGACTTTGCGCAC 4843
Oy 241 PheMetTyrGluLeuLeuCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGlu 260
Db 4844 TTCAATGATGAACCTGTGTGGAGAGCAGACCCCTTCTCTATCCAGGCGCAGAG 4903
Oy 261 TyrThrValAlaPheTrpAspGlyLysGluGluSerArgAsnTyrThrValAlaAsnThr 280
Db 4904 TACACCGTGTGTGTTTGGATGGAAGAAAGAGATCCAGAGACTACAGCGTGTGAACACA 4963
Oy 281 GluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
Db 4964 GAGAAAGGCGCTCATGAGAGGTGCAGAGATGTGTGCTGCCCTGGAGTGAAGGTAGCTGCCAG 5023
Oy 301 LeuGlnValGlnArgSerLeuTrpThrAlaThrGlnAspGlnLysIleTyrIleTyrThr 320
Db 5024 CTCAGGTCCAGAGATCCCTGTGACACCCAGCAGG----- 5059
Oy 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAla 340
Db 5059 ----- 5059
Oy 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
Db 5060 -----AATTCCTACCTGTGCTGTAGCGGCGCTC 5086
Oy 361 AlaAspGlyLeuValAlaValPheProValAlaArgGlyThrProLysAspSerCysSer 380
Db 5087 GCCATGGGCTGTGTGCTGTGTCTCCGTGTGCGGGCCACCCCAAGAGACGCTGTCTCC 5146
Oy 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
Db 5147 TACCTGTCTCTACACACAGCCACAGAGTCCAAAGTTCACATCCGCGAGTGAAGAGCAGG 5206
Oy 401 GlnAsnProTyrProValLysAlaMetGluValAlaAsnSerGlySerGluValTyrPyr 420
Db 5207 CAGAACCCCTACCCAGTGAAGGCCATGAGAGGTGTCAACAGCGGCTGTGAGGTCTGTGATC 5266
Oy 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
Db 5267 AGCAATGGAGCGGCGCTCTCTGTCTCATGACATGTGCTCTCTGGAATGTGCAGGGCGCTG 5326
Oy 441 GluProTyrMetAlaProSerMetValThrSerValAlaCysSerSerGluGlyArgGly 460
Db 5327 GAGCCCTACATGCGCCCTCCATAGTTAGTACGTACGTGAGTGTGAGGAGGAGAGG 5386
Oy 461 GluGluValAlaTyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
Db 5387 GAGGAGGTGCTGTGTGTGTGATGACAAAGGCAACTCTCTGTGTATGTACACATCCACC 5446
Oy 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
Db 5447 ACCTACCACTGTGTGTGCGGCTGTGTGCGGGGTCCCGAGCCCTCAGGAGCATGTGTT 5506
Oy 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
Db 5507 CCGGTGCGGCGCTTGGAGACGGAACCCCGGAGCAGACAGCCCAACCAAGAGTG 5566
Oy 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluGluGlnGlyThrGln 540
Db 5567 CCGTGAAGGGGACTCATCGCGAGCTGAGCATCATGATGAGAGAGCTGGGACGCGAG 5626

```


Pred. No.: 0 Length: 3496
 Score: 3822.00 Matches: 728
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-836-392-21 (1-728) x US-09-836-392-8 (1-3496)

QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLeuLysAsnPheserGluPheArgGlnGlu 20
 DB 22 ATGCTGAGGACCTGGCGGCGACCGATGCGCATGAAGAACTTCTCCGAGTTCGGGACGAG 81
 QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
 DB 82 GCGAGATGCTGCACGCGCGTCGACACCCCTGCATGTGGCGCTCATCGGCATCAGATC 141
 QY 41 HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 DB 142 CACCCGCTGCTTGGCCCTGGAGCTCGCGCCCTCAGCGCTCAACACCGCTGTGTC 201
 QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnLysIleAla 80
 DB 202 GAGAAAGCCAGAGATCTCTCTTATACCCCTGGGACACATGCTCACCCCAAAATAGCC 261
 QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
 DB 262 TACCAAGTGCCTCGGCGCTCGGCTACCTGCACAGAAAACATCATCTCTGTGACCTG 321
 QY 101 LysSerAspAsnIleLeuValTyrPheSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
 DB 322 AAGCGGACAAACATTTCTGGTGTGCTCCCTGAGCTCAAGAGACACATCAACATCAAGCTA 381
 QY 121 SerAspTyrGlyIleSerArgGlnSerPheHisGluGlyAlaLeuGlyValGluGlyThr 140
 DB 382 TCTGACTACGGGATTTGAGAGCAGCTATTCATGAGGGCCCTAGGCGTGGAGGGACT 441
 QY 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGlyLysValAspMet 160
 DB 442 CTGGGCTACAGAGGCCAGAGATCAGGCGCTCGCATGTGTATGATGAGAGATGATGATG 501
 QY 161 PheSerTyrGlyMetValLeuTyrGlyLeuLeuSerGlyGlnArgProAlaLeuGlyHis 180
 DB 502 TTCTCTCATGTGAATGTGTCTCTCAGAGTTCTCTCAGAGACGCGCTGCACCTGGGCGAC 561
 QY 181 HisGlnLeuGlnIleAlaLysLysLeuSerTyrGlyTyrLeuArgProValLeuGlyGlnPro 200
 DB 562 CACCAAGCTCCAGATGCGCAAGAGCTGTCCAGAGGCTATCGCGGCTTCTGGGCGACCGC 621
 QY 201 GluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGluCysTyrPheAspThrLysPro 220
 DB 622 GAGGAAGTGCAGTTCGCGGCGACCTGCAGCGCTCATGTAGAGTGTGGGACATAGACCA 681
 QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
 DB 682 GAAAGAGGACCGGCTGGCGCTGTGTGTGTGAGCCAGATGAAGACCGGACTTTTGGCCACC 741
 QY 241 PheMetTyrGluLeuCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGlu 260
 DB 742 TTCTATGATATACCTGTCTGTGGGAACAGACCTTCTTCATCTCCAGGGCCAGAG 801
 QY 261 TyrThrValValPheTyrPheAspGlyLysGluLysArgAsnTyrThrValAlaAsnThr 280
 DB 802 TACACCGTGTGTTTGGATGGAAAAGAGAGATCCAGAACTACACGAGTGTGTGAACACA 861
 QY 281 GluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
 DB 862 GAGAAAGGCGCTCATGAGAGTGCAGAGAGATGTGCTGCGCTGGATGAAGGTGAGGTGCAG 921
 QY 301 LeuGlnValAlaLysSerLeuTyrThrAlaThrGluAspGlnLysIleTyrIleTyrThr 320
 DB 922 CTCACAGTCCAGAGATCCCTGTGTGAGCACCGAGAGACAGAAAATCTACATCTACAC 981

QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
 DB 982 CTCAGAGGATGTGGCCCTTAAACACACCCAGCCCTGGATATCTCAGCTGTGCTG 1041
 QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
 DB 1042 ACCTGCTTGTGGCCGCTGTATTAAGAAATTCCTACCTGTGTATAGCGGCGCTC 1101
 QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 DB 1102 GCGATGTGGCTGTGGCTGTGTCCGTGTGTGGGGACCCCAAGAGACGCTGTCTC 1161
 QY 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
 DB 1162 TACCTGTGTACACACACAGCAAGCTCAAGTTAGCATGCGGATGAAGCGCAGG 1221
 QY 401 GluAsnProTyrProValLysAlaMetGluValAlaAsnSerGlySerGluValTyrTyr 420
 DB 1222 CAGAACCCCTTACCAGATGAAGGCGCATGAGTGTGCACAGCGGCTGAGGCTGTGAC 1281
 QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
 DB 1282 AGCATGTGGCGGCGCTCTGTGTATGAGTGTGTCTCCCTGGAGATGTGACGGCGCTG 1341
 QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGluArgGly 460
 DB 1342 GAGCCCTACATGGCCCGCTCCATAGTTAGTCACTGCTGTGTGACGCTGTAGGGCAGAGG 1401
 QY 461 GluGluValValTyrPheLysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 DB 1402 GAGAGAGT 461
 QY 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
 DB 1462 ACCTACAGCTGT 1521
 QY 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
 DB 1522 CCGGTGCGGCGCTGTGACACAGGAACCCCGGCGACGACGACGACGACGACGACGACG 1581
 QY 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluLeuGlyThrGln 540
 DB 1582 CCGAGAGGAGACTCCATCGCGGCGAGCTGAGCATCATGTATGAGTGTGGGCGACGAG 1641
 QY 541 IleLeuIleHisGlnLeuSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSer 560
 DB 1642 ATCTGATCCACAGGAATCAGTCACTGACTGTCTCATGTCTCTCATCTCATCTC 1701
 QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
 DB 1702 CCACCCCGGCGAGCTGCCAGGTCCCTCAAGCTCCCGCAGCTCCCGCAGCAAGTTCTTCC 1761
 QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
 DB 1762 AGTGTCTCTTCCACACCACTGCGAGACTCAGACATGCTACATAGCCCGGCTGTGTC 1821
 QY 601 SerAspArgSerGluHisAspLeuThrPrometAspGlyGluThrPheSerGlnHisLeu 620
 DB 1822 TCCGAGAGGTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1881
 QY 621 GlnAlaValLysIleLeuAlaValAlaArgAspLeuIleTyrAlaProArgArgGlyLysAsp 640
 DB 1882 CAGCGCGTGAAGTCTCTCCCGCTCAAGACCTCATTTGGTCCCGCAGCGCGGCTGGAAGT 1941
 QY 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 DB 1942 GTTATGCTATGTGGCTGTGAAGAGATGTGAAGCCAGCGGGGCGGAGTATGTGCGCTG 2001
 QY 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 680
 DB 2002 TTAAGAGCCCGAGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2061
 QY 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrCysLeuAlaValTyrArg 700

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 15, 2003, 08:37:24 ; Search time 123 Seconds

(without alignments)
5191.692 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822
Sequence: 1 MLRHLRATDAKNFSEFRQE.....IFYQSYELGRLEACTRRKR 728

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame.p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09836392/runat_08042003_090542_23517/app_query.fasta_1.903
-DB=PublishedApplications.NA -QEXT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=71 -MATRIX=blosum62
-TRANS=human0.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTEXT=pct -NOR=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09836392_@CGN_1.1_69_etunat_08042003_090542_23517
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications.NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3822	100.0	3496	9	US-09-836-392-8
2	3579.5	93.7	7015	9	US-10-132-382-5
3	3579.5	93.7	7093	9	US-10-132-382-1
4	3579.5	93.7	7229	9	US-10-132-382-7

5	3579.5	93.7	7307	9	US-10-132-382-3	Sequence 3, Appl1
6	307.5	8.0	3454	10	US-09-969-347-226	Sequence 226, App
7	270.5	7.1	1662	9	US-09-938-842A-1014	Sequence 1014, App
8	269.5	7.1	3111	12	US-10-014-882-1	Sequence 1, Appl1
9	269.5	7.1	3518	12	US-10-014-882-3	Sequence 3, Appl1
10	268	7.0	1428	9	US-09-938-842A-882	Sequence 882, App
11	260.5	6.8	3558	9	US-10-143-133-1	Sequence 1, Appl1
12	259.5	6.8	1638	9	US-09-938-842A-903	Sequence 903, Appl
13	251	6.6	2211	9	US-09-938-842A-1577	Sequence 1577, App
14	247	6.5	2770	9	US-09-977-260-5	Sequence 5, Appl1
15	247	6.5	2770	9	US-09-977-261-5	Sequence 5, Appl1
16	247	6.5	2770	10	US-09-977-269-5	Sequence 5, Appl1
17	243.5	6.4	2863	10	US-09-934-842A-1631	Sequence 1631, App
18	243.5	6.4	7607	10	US-09-982-610-19	Sequence 1, Appl1
19	243	6.4	1365	10	US-09-757-982-6	Sequence 4, Appl1
20	243	6.4	2120	10	US-09-757-982-4	Sequence 1270, App
21	242.5	6.3	2640	9	US-09-938-842A-1270	Sequence 1, Appl1
22	242.5	6.3	3286	10	US-09-904-389-1	Sequence 1, Appl1
23	239	6.3	3270	10	US-09-835-081-1	Sequence 1190, Appl
24	238.5	6.2	1239	9	US-09-938-842A-1190	Sequence 128, App
25	238	6.2	3045	8	US-08-910-386A-16	Sequence 1354, App
26	238	6.2	3124	9	US-09-974-298-138	Sequence 16, App
27	238	6.2	3358	9	US-09-954-531-1554	Sequence 1354, App
28	237.5	6.2	2016	10	US-09-758-003-1	Sequence 1, Appl1
29	237	6.2	1627	10	US-09-828-313-16	Sequence 16, Appl
30	232	6.1	772	10	US-09-770-445-908	Sequence 908, App
31	230	6.0	2124	9	US-09-938-842A-1079	Sequence 1079, App
32	229	6.0	5180	10	US-09-954-475-1246	Sequence 1246, App
33	229	6.0	5180	10	US-09-880-107-3668	Sequence 3668, App
34	228.5	6.0	1788	9	US-10-158-895-14	Sequence 14, Appl1
35	228.5	6.0	2656	9	US-10-158-895-3	Sequence 3, Appl1
36	227	5.9	2651	9	US-09-938-842A-1085	Sequence 1085, App
37	226.5	5.9	1179	9	US-09-938-842A-1559	Sequence 342, App
38	225	5.9	1308	9	US-09-938-842A-1559	Sequence 1559, App
39	225	5.9	1367	10	US-09-925-302-262	Sequence 262, App
40	225	5.9	2637	12	US-10-060-332-1	Sequence 1, Appl1
41	225	5.9	2975	9	US-09-991-417-25	Sequence 25, Appl1
42	225	5.9	2962	9	US-09-991-417-106	Sequence 106, Appl1
43	222.5	5.8	2437	9	US-09-158-723-3	Sequence 3, Appl1
44	222.5	5.8	3503	10	US-09-823-038A-58	Sequence 58, Appl1
45	221.5	5.8	1170	9	US-09-938-842A-647	Sequence 647, App

ALIGNMENTS

RESULT 1
US-09-836-392-8
Sequence 8, Application US/09836392
Patent No. US20020173458A1
GENERAL INFORMATION:
APPLICANT: Ruden et al.
TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptide
FILE REFERENCE: PTO20P1
CURRENT APPLICATION NUMBER: US/09/836,392
CURRENT FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: PCT/US00/28066
PRIOR FILING DATE: 2000-10-11
PRIOR APPLICATION NUMBER: 60/159,542
PRIOR FILING DATE: 1999-10-15
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: 60/189,027
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 3496
TYPE: DNA
ORGANISM: Homo sapiens
US-09-836-392-8
Alignment Scores:

```

QY 145 AlaProgluilearProargileValtyrAspGluLysValAspMetPheSerTyrGly 164
DB 520 GCTCAGAGATGATCCAGAGAGCTCCCGTGCAGAAACTGTGACACATATCTCTATGCT 579
QY 165 MetValLeuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlnHisGlnLeuGln 184
DB 580 GTGCTTCTCTGGAGATGCTTAACAGAGAGGCTCCCTTTAAAGTTTGAAGATTTACAA 639
QY 185 IleAla---LysLysLeuSerLysGlyIleArgProValLeuGlnProGluGluVal 203
DB 640 GTACCTTGCGCTTGAGTGGAAAAAAGCAGAGA-----TTAACCAATTCAGAGCTTGC 693
QY 204 GlnPheArgArgLeuGlnAlaLeuMetMetGluCysTrpAspThrLysProGluLysArg 223
DB 694 CCC---AGAACTTTGCTGCACTGTACATCAGTGTGGAGAGCTGATGCCAGAAACGG 750
QY 224 ProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThrPheMetTyr 243
DB 751 CCATCAATTCAGCAAAATCATTCATCTCGAGTCCATGTCATGAATGACAGAGCTTCCT 810
QY 244 GluLeuGly 246
DB 811 GACAGAGTGT 819

RESULT 15
US-09-163-115-6
Sequence 6, Application US/09163115A
Patent No. 6183962
GENERAL INFORMATION:
APPLICANT: Action, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/163,115A
CURRENT FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 1365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1365)
US-09-163-115-6

Alignment Scores:
Pred. No.: 3,86e-14 Length: 1365
Score: 243.00 Matches: 71
Percent Similarity: 49.79% Conservative: 50
Best Local Similarity: 29.22% Mismatches: 104
Query Match: 6,36% Indels: 18
DB: 4 Gaps: 8

US-09-836-392-21 (1-728) x US-09-163-115-6 (1-1365)
QY 10 AlaMetLysAsnPheserGluPheArgGlnGlnAlaSerMetLeuHisAlaLeuGlnHis 29
DB 127 GCTGTAAAGAGCTCTCCCAAAATAGAGAAAGAGCAAAATACCTGCTGCTCAGTCTAC 186
QY 30 ProCysIleValAlaLeuLeuGlyIleSerIleHisProLeuGlyPheAlaLeu----- 47
DB 187 AGAAACATCAATCCAGTTTATGAGATATCTGTGACCTCCCACTATGCGATGTGACA 246
QY 48 GluLeuAlaProLeuSerSerLeuAsnThrValLeuSerGluAsnAlaArgAspSerSer 67
DB 247 GAATATGCTTCCTGAGTACTATGATTACATTACAGTAACAGCA-----AGTGAG 300
QY 68 PheIleProLeuGlnHisMetLeuThrGlnLysIleAlaTyrGlnIleAlaSerGlyLeu 87
DB 301 GAGATGATATGATGATCATTATGACCTCG-----GCCACTGATGTAGCCAAAGGATG 354
QY 88 AlaTyrLeuHisLysLys-----AsnIleIlePheCysAspLeuLysSerAspAsn 104

```

```

DB 355 CATTATTACATATGAGAGGCTCCTGTCAAGGTGATTCACAGAGAGCCCAAGTCAAGAAAC 414
QY 105 IleLeuValTrpSerLeuAspValLysGlnHisIleAsnIleLysLeuSerAspTyrGly 124
DB 415 GTTGTATTACCTGCTGATGGAGTA-----CTGAAGATCTGTGACTTTGGT 459
QY 125 IleSerArgIleSerPheHisGlnGlyAlaLeuGlnValAlaGluGlyThrProGlyTyrGln 144
DB 460 GCCTCTGCTGCCATTAACCATACACACACATGCTCTGTTGGAGACTTCCCATGGATG 519
QY 145 AlaProgluilearProargileValtyrAspGluLysValAspMetPheSerTyrGly 164
DB 520 GCTCAGAGATGATCCAGAGAGCTCCCGTGCAGAAACTGTGACACATATCTCTATGCT 579
QY 165 MetValLeuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlnHisGlnLeuGln 184
DB 580 GTGCTTCTCTGGAGATGCTTAACAGAGAGGCTCCCTTTAAAGTTTGAAGATTTACAA 639
QY 185 IleAla---LysLysLeuSerLysGlyIleArgProValLeuGlnProGluGluVal 203
DB 640 GTACCTTGCGCTTGAGTGGAAAAAAGCAGAGA-----TTAACCAATTCAGAGCTTGC 693
QY 204 GlnPheArgArgLeuGlnAlaLeuMetMetGluCysTrpAspThrLysProGluLysArg 223
DB 694 CCC---AGAACTTTGCTGCACTGTACATCAGTGTGGAGAGCTGATGCCAGAAACGG 750
QY 224 ProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThrPheMetTyr 243
DB 751 CCATCAATTCAGCAAAATCATTCATCTCGAGTCCATGTCATGAATGACAGAGCTTCCT 810
QY 244 GluLeuGly 246
DB 811 GACAGAGTGT 819

```

Search completed: April 15, 2003, 08:59:00
 Job time : 84 secs


```

Db 301 GAGATGATATGATCATATTATGACCTGG-----GCCACTGATGAGCCAAAGGATG 354
QY 88 AATATLeuHisLysLys-----AsnIlellePheCysAspLeuLysSerAspAsn 104
Db 355 CATTATTTTACATATGAGAGGCTCTGTCAAGGATATTCACAGAACCTCAAGAAC 414
QY 105 IleLeuValTrpSerLeuAspValLysGluHisIleAsnIleLysLeuSerAspTyrGly 124
Db 415 GTTGTTATAGCTCTGATGAGTA-----CTGAAGATCTGTGACTTTGGT 459
QY 125 IleSerArgGlnSerPheHisGluGlyAlaLeuGlyValGluGlyThrProGlyTyrGln 144
Db 460 GCCTCTGCGTTCCATTAACCATACACACATGTCCTGTGGTGGTCCATTCCTCATGGATG 519
QY 145 AlaProGluIleArgProArgIleValTyrAspGluLysValAspMetPheSerTyrGly 164
Db 520 GCTCCAGAAAGTTATCCAGAGCTCCCTGTGTCAAAAATTGTACACATATTCCTAAGGT 579
QY 165 MetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHisGlnLeuGln 184
Db 580 GTGTTCTCTCGGAGATGCTTACACAGGAGGTCCTTTAAAGTTTGGAAAGATTTCACA 639
QY 185 IleAla--LysLysLeuSerLysGlyIleArgProValLeuGlnProGlyGluVal 203
Db 640 GTAAGTTGGCTGTAGTGAAGAAAAACAGAGAGA-----TTAACCATTCACAGAGCTTGC 693
QY 204 GlnPheArgArgLeuGlnAlaLeuMetGluCysTrpAspThrPheAlaThrPheMetTyr 223
Db 694 CCC--AGAAATTGCTGTGACGTGTACATCAGTGTGGAGAGCTGATGCCAAGAAACGG 750
QY 224 ProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThrPheMetTyr 243
Db 751 CCATCATTCACAGAAATCATTTCAATCTGAGTCCATGTCATTAATGACAGAGCTTCT 810
QY 244 GluLeuCys 246
Db 811 GACACAGTGT 819

RESULT 10
US-09-221-928-6
; Sequence 6, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAR-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMT-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-928-6

Alignment Scores:
Pred. No.: 3 86e-14 Length: 1365
Score: 243.00 Matches: 71
Percent Similarity: 49.79% Conservative: 50
Best Local Similarity: 29.22% Mismatches: 104
Query Match: 6.36% Indels: 18
DB: 3 Gaps: 8

US-09-836-392-21 (1-728) x US-09-221-928-6 (1-1365)
QY 10 AlaMetLysAsnPheSerGlnPheArgGlnGluAlaSerMetLeuHisAlaLeuGlnHis 29
||||:||||| :|||:||||| :|||:||||| :|||:|||||

```

```

Db 127 GCTTAAGAAAGCTCCTCAAAATAGAGAAAGGACAAATATCTAGTCTCACTAC 186
QY 30 ProCysIleValAlaLeuIleGlyIleSerIleHisProLeuSerPheAlaLeu----- 47
Db 187 AGAAACATATCCAGTTTATGAGTAATTTTGAACTCCCACTATGCGATTGTCACA 246
QY 48 GluLeuAlaProLeuSerSerLeuAsnThrValLeuSerGluAsnAlaArgAspSer 67
Db 247 GAAATAGCTTCTCTGGGATCAGCTATGATTAATTAACAGTAACAGTA-----AGTAG 300
QY 68 PheIleProLeuGlnHisMetLeuThrGlnLysIleAlaTyrGlnIleAlaSerGlyLeu 87
Db 301 GAGATGATATGATGATCATATTATGACCTGG-----GCCACTGATGAGCCAAAGGATG 354
QY 88 AATATLeuHisLysLys-----AsnIlellePheCysAspLeuLysSerAspAsn 104
Db 355 CATTATTTTACATATGAGAGGCTCTGTCAAGGATATTCACAGAACCTCAAGAAC 414
QY 105 IleLeuValTrpSerLeuAspValLysGluHisIleAsnIleLysLeuSerAspTyrGly 124
Db 415 GTTGTTATAGCTCTGATGAGTA-----CTGAAGATCTGTGACTTTGGT 459
QY 125 IleSerArgGlnSerPheHisGluGlyAlaLeuGlyValGluGlyThrProGlyTyrGln 144
Db 460 GCCTCTGCGTTCCATTAACCATACACACATGTCCTGTGGTGGTCCATTCCTCATGGATG 519
QY 145 AlaProGluIleArgProArgIleValTyrAspGluLysValAspMetPheSerTyrGly 164
Db 520 GCTCCAGAAAGTTATCCAGAGCTCCCTGTGTCAAAAATTGTACACATATTCCTAAGGT 579
QY 165 MetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHisGlnLeuGln 184
Db 580 GTGTTCTCTCGGAGATGCTTACACAGGAGGTCCTTTAAAGTTTGGAAAGATTTCACA 639
QY 185 IleAla--LysLysLeuSerLysGlyIleArgProValLeuGlnProGlyGluVal 203
Db 640 GTAAGTTGGCTGTAGTGAAGAAAAACAGAGAGA-----TTAACCATTCACAGAGCTTGC 693
QY 204 GlnPheArgArgLeuGlnAlaLeuMetGluCysTrpAspThrPheAlaThrPheMetTyr 223
Db 694 CCC--AGAAATTGCTGTGACGTGTACATCAGTGTGGAGAGCTGATGCCAAGAAACGG 750
QY 224 ProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThrPheMetTyr 243
Db 751 CCATCATTCACAGAAATCATTTCAATCTGAGTCCATGTCATTAATGACAGAGCTTCT 810
QY 244 GluLeuCys 246
Db 811 GACACAGTGT 819

RESULT 11
US-09-221-527-6
; Sequence 6, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAR-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMT-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-527-6

```

FILING DATE: 04-APR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Wendy M. Lee
 REGISTRATION NUMBER: 00,000
 REFERENCE/DOCKET NUMBER: 821P3PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/952-9881
 TELEFAX: 415/952-9881
 TELEEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7607 bases
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 PCT-US95-04228-19

Alignment Scores:
 Pred. No.: 5.65e-13 Length: 7607/
 Score: 243.50 Matches: 99
 Percent Similarity: 45.38% Conservative: 63
 Best Local Similarity: 27.73% Mismatches: 117
 Query Match: 6.37% Indels: 79
 DB: Gaps: 16

US-09-836-392-21 (1-728) x PCT-US95-04228-19 (1-7607)

QY 15 SerGluPheArgGlnGluAlaSerMetLeuHisAlaLeuGlnHisProGlyIleValAla 34
 DB 2562 AATGACTCTCTGAGGGGACAGATATATAGAACTTATTTACAGAGTTGAGAGCAT 2503
 QY 35 LeuIleGlyIleSerIle-----HisProLeuGlyPheAlaLeuGlnValAlaProLeu 52
 DB 2502 CTTTATGCTTTTGACATTAGAAATCAATTTATTTTACAGAGTTGAGAGCAT 2443
 QY 53 SerSerLeuAsnThrValLeuSerGluAsnAlaArgSerPheIleProLeuGly 72
 DB 2442 GGAAGTCTCGAAGAAATATCTC-----CAAAATGACACGTGATCAAAATTC----- 2398
 QY 73 HisMetLeuThrGlnGlyIle-----AlaTyrGlnIleAlaSerGlyLeuAlaTyrLeu 90
 DB 2397 ---CATGTGCTCAACAGGTGACAGATGCGCACAGGTCCTGGAATGGCTATCTG 2341
 QY 91 HisLeuAsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValTyrSerLeu 110
 DB 2340 GAGTCTCGAATACATCATTCACAGATCTGCTGCCGAATGTCTC----- 2293
 QY 111 AspValIleGlnHisIleAsnIleLysLeuSerAspTyrGlyIleSerArg----- 127
 DB 2292 ---GTTGGTGAACATATATCTACAAAGTACAGATTTTGCACTTGCCACAGTTTAAAG 2236
 QY 128 -----GlnSerPheHisGlnGlyAlaLeuGlyValAlaGlyTyr 140
 DB 2235 GTAGATAATGACACATCTATGATCTAGACAGAAATATAAGTCGCGGAGTGGACT 2176
 QY 141 ProGlyTyrGlnAlaProGluIleArgProGlyIleValTyrAspGluLysValAspMet 160
 DB 2175 -----GCGCCGGAAGCCATTCTGTAATTAATTCAGCATTAAGTCCGAGTA 2128
 QY 161 PheSerTyrGlyMetValLeuTyrGluLeuLeuSer---GlyGlnArgProAlaLeuGly 179
 DB 2127 TGGTCAATTTGGAATCCTCTTTATGAATATCTACTATGCGAAATGCTTACAGAGGT 2068
 QY 180 HisHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGln 199
 DB 2067 ATGACAGTGTCCAGGAAATCCAGATGTTGCTCAAAACATATAGA-----CTTCCGCAA 2014
 QY 200 ProGluGlnValGlnPheArgArgLeuGlnAlaLeuMetGluCysTyrPaspThrLys 219
 DB 2013 CCATCCAACTGTGCA---CAGCAATTTTACACATCATGTTGAGTCTGGAATGCAGAG 1957
 QY 220 ProGluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAla 239
 |||:|||||

DB 1956 CTAAGAACGA-----CCTACATTGAG 1933
 QY 240 ThrPheMetIyrProLeuGlyCysGlyGlnThrAlaPhe----- 254
 DB 1932 ACACGTGGCTTGGAACTGTAAGACTATTTTGAACAGA-CTCTCATATTACAGATGCAAA 1874
 QY 255 -----SerSerGlnGlyGlnIleTyrThrValAlaPheTyrPaspGlyLysGlu 270
 DB 1873 TAATCTCATAGATGAACACTGAGAGAAAGATATCAATATATAAGTACAAACAAATTC 1814
 QY 271 Glu-----SerArgAsnTyrThrValValAlaSerThrGluLysGlyLeuValGln 288
 DB 1813 AATTAATCCATTCCAAAATATCATATGATTCAC----- 1781
 QY 289 ArgMetCysCysProGlyMetLysValSerCysGlnLeuGlnValGlnArgSerLeuTyr 308
 DB 1780 -----CACTGCAACATCATGATTATCTGACA 1754
 QY 309 ThrAlaThrGluAspGlnLysIleTyrIleTyrThrLeuLysGly---MetCysProLeu 327
 DB 1753 TATTCAGATGATGATGAATGAAGTGGCCATGTATATGAAAATATATTGTCGATTTA 1694
 QY 328 AsnThrProGlnGlnAlaLeuAspThrProAlaValAlaThrCysPheLeu 344
 DB 1693 TTGACTGGCGAACACTGC---AGACAGATCAAGGTATATATATTTCTC 1646

RESULT 9
 US-09-221-235-6
 ; Sequence 6, Application US/09221235
 ; Patent No. 6043040
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: INT-050
 ; CURRENT APPLICATION NUMBER: US/09/221, 235
 ; EARLIER FILING DATE: 1998-12-28
 ; EARLIER APPLICATION NUMBER: 09/163, 115
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 1365
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1365)
 ; US-09-221-235-6

Alignment Scores:
 Pred. No.: 3.86e-14 Length: 1365
 Score: 243.00 Matches: 71
 Percent Similarity: 49.79% Conservative: 50
 Best Local Similarity: 29.22% Mismatches: 104
 Query Match: 6.36% Indels: 18
 DB: Gaps: 8

US-09-836-392-21 (1-728) x US-09-221-235-6 (1-1365)

QY 10 AlaMetIysAsnPheSerGluPheArgGlnGluAlaSerMetLeuHisAlaLeuGlnHis 29
 DB 127 GCTGTAAAGAGCTCTCTCAAAATAGAGAAAGGACGAAATATCTAGCTGCTCACTCAC 186
 QY 30 ProCysIleValAlaLeuIleGlyIleSerIleHisProLeuGlyPheAlaLeu----- 47
 DB 187 AGAACAATATCATCAAGTTTATGAGATTAATCTGAACCTCCACATATGACATTCACACA 246
 QY 48 GluLeuAlaProLeuSerSerLeuAsnThrValLeuSerGluAsnAlaArgAspSerSer 67
 DB 247 GAATATGCTTCTCTGGAGACACCTATGATTAATTAACATTAACAGTACAGA-----ACTGAG 300
 QY 68 PheIleProLeuGlyHisMetLeuThrGlnLysIleAlaTyrGlnIleAlaSerGlyLeu 87
 : : : : : |||:|||||

```

STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,648
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P082193PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ. ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 7607 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-446-648-19

Alignment Scores:
Pred. No.: 5.65e-13 Length: 7607
Score: 243.50 Matches: 99
Percent Similarity: 45.38% Conservative: 63
Best Local Similarity: 27.73% Mismatches: 117
Query Match: 6.37% Indels: 79
Gaps: 16
DB: 4

US-09-836-392-21 (1-728) x US-08-446-648-19 (1-7607)
QY 15 SerGIuPhaRgInGIuAlaSerMetLeuHISAlaLeuGlnHISProCysIleValAla 34
DB 2562 AATGACTCTCTGAGGCGACAGATCTAGAGAACTTAGACATCCAAAGCTTATCCAG 2503
QY 35 LeuIleGIyIleSerIle-----HISProLeuGlySerPheIleProLeuGly 52
DB 2502 CTTTATGCTGTTGCACTTGAAGATCAATTATTTATTACAGAGTTGATGAGACAT 2443
QY 53 SerSerLeuAsnThrValLeuSerGIuAsnAlaArgAspSerSerPheIleProLeuGly 72
DB 2442 GGAAGCTCGAAGAAATATCTC---CAAAATGACACTGATCAAAAATC----- 2398
QY 73 HIsMetLeuThrGlnIleSile-----AlaTyrGlnIleAlaSerGIyLeuAlaTyrLeu 90
DB 2397 ---CATCTGACTCAACGCTAGACATGCGGACAGCTTGTCTGGAATGGCCTATCTG 2341
QY 91 HIsLysLysAsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValIlePheSerLeu 110
DB 2340 GAGTCTCGAAGTACATTCACAGAGATGCTGCTCCAGAAATGTCTC----- 2293
QY 111 AspValLysGIuHISLysIleAsnIleLysLeuSerAspTyrGIyIleSerArg----- 127
DB 2292 ---GTGCTGACATATATATCTCAAAAGTAGACATTTTGGACTTGCACAGAGTTTAAAG 2236
QY 128 -----GlnSerPheHISGluLysIleValAlaLeuGIyValGIyLysThr 140
DB 2235 GTAGATATAGAGACATCTATGATCTAGACAGAAATTAAGCTCCGCTGAAGTGAAGTACT 2176
QY 141 ProGIyLysGIuAlaProGIuLysIleArgProArgIleValIleTyrAspGIuLysValAspMet 160

```

```

DB 2175 -----GCGCCGGAAGCCATTGTAATTAATTCAGATTAGTCCGATGTA 2128
QY 161 PheSerTyrGIyMetValLeuTyrGIuLeuLeuSer---GlyGIuArgProAlaLeuGIy 179
DB 2127 TGGTCATTTGGAAATCCTTTCTTATGAATCAATTATATGCGAAATGCCATTACACTGGT 2068
QY 180 HIsHISGIuLeuGlnIleAlaLysLysLeuSerGIyLysArgProValLeuGIyGln 199
DB 2067 ATACAGAGTCCAGGATTAATCCAGATGTGGCTCAAAACATATGA-----CTTCCGCA 2014
QY 200 ProGIuGIuValGIuPheArgArgLeuGlnAlaLeuMetLeuGIyCysTyrAspThrLys 219
DB 2013 CCATCCAACTGTCCA---CAGCAATTTTACAAATCATCATGTGGAGCTGGATGAGAG 1957
QY 220 ProGIuLysArgProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAla 239
DB 1956 CCTAAGAACGA-----CCTACATTTGAG 1933
QY 240 ThrPheMetTyrGIuLeuCysGlyLysGlnThrAlaPhePhe----- 254
DB 1932 ACAGTGGCTGGAAACTTGAAGACTATTTGAACAGCA-CTCTCATATTCAGATGCAAA 1874
QY 255 -----SerSerGIuGIyGlnGIuTyrThrValValPheThrAspGIyLysGIu 270
DB 1873 TAAGTTCATATAGATGAAGACTGAGAGAAATATCAATATATAAGTAGCAAAACAATTC 1814
QY 271 Glu-----SerArgAsnTyrThrValValAsnThrGIuLysGIyLeuMetGIuValGln 288
DB 1813 AAATATTCATTCACAAATTCAAATGTTATTCAC----- 1781
QY 289 ArgMetCysCysProGIyMetLysValSerCysGIuLeuGlnValGlnArgSerLeuTyr 308
DB 1780 -----CACTGCACATCACTTATTCCTGACA 1754
QY 309 ThrAlaThrGIuAspGIuLysIleTyrIleTyrThrLeuLysGIy---MetCysProLeu 327
DB 1753 TATTCAGATATAGATGAAGAGTTGGCCATATATTAAGAAAGATATTTATGTCATTTTA 1694
QY 328 AsnThrProGIuGlnAlaLeuAspThrProAlaValAlaThrCysPheLeu 344
DB 1693 TTGACTGGCAACACTGC---AGGACAGTCAAGGTGATATATATTTCTTCCCTC 1646

RESULT 8
PCT-US95-04228-19/C
Sequence 19, Application PC/TUS9504228
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616

```


QY 91 HisLysLysAsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValTrpSerLeu 110
 Db 2340 GAGTCTCGGAACTACATTCTCAGAGATCTCGCTCCAGAAATGCTCTC----- 2233
 QY 111 AspValLysGluHisIleAsnIleLysLeuSerAspTyrGlyIleSerArg----- 127
 Db 2292 ---GTTGGTGAACATATAATACATCAAAAGTAGAGAGATTTTGGACTTGCAGAGATTTTAAAG 2236
 QY 128 -----GlnSerPheHisGluGlyValLeuGlyValGluGlyThr 140
 Db 2235 GTAGATTAATGAGACATCTATGATGATTTAGACACGAAATTAAGCTCCCGGTGAAGTGACT 2176
 QY 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluValValAspMet 160
 Db 2175 -----GCGCCGGAAGGCATTCGTAGTAATTAATTCAGCATTAAGTCGATGTA 2128
 QY 161 PheSerTyrGlyMetValLeuTyrGluLeuLeuSer--GlyGlnArgProAlaLeuGly 179
 Db 2127 TGGCATTTTGGATCCCTCTTTTGAAGAACTACTTACTTAGCGCAAAATGCTTCACGTGT 2068
 QY 180 HisHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGln 199
 Db 2067 ATGACAGGTGCCAGGATATCCAGATGTTGGCTCAAAACTATAGA-----CTTCCGCA 2014
 QY 200 ProGluGluValGlnPheArgArgLeuGlnAlaLeuMetGluCysTyrTrpAspThrLys 219
 Db 2013 CCATCCAAACGTGCA---CAGCAATTTTACACATCATGTTGGAGTGGAGATCAAG 1957
 QY 220 ProGluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAla 239
 Db 1956 CCTAGAGCAACGA-----CCTACATTGAG 1933
 QY 240 ThrPheMetTyrGluLeuCysCysGlyLysGlnThrAlaPhe----- 254
 Db 1932 ACACTGCGTTGGAACCTGAAAGCTATTTGAAACGA-CTCTCATATTCAGATGCAA 1874
 QY 255 -----SerSerGlnGlyGlnGluTyrThrValValPheTrpAspGlyLysGlu 270
 Db 1873 TAACTTCATTAAGATGAACACTGGAGAGAAATATCAATATTAAGACAAACAAATTC 1814
 QY 271 Glu-----SerArgAsnTyrThrValValAsnThrGluGlyLeuMetGluValGln 288
 Db 1813 AAAAATATCCATTCCAAATAATCAATGTTATCAAC----- 1781
 QY 289 ArgMetCysCysProGlyMetLysValSerCysGlnLeuGlnValGlnArgSerLeuTrp 308
 Db 1780 -----CAACTGCACAAATCAGTTATCTGTACA 1754
 QY 309 ThrAlaThrGluAspGluLysIleTyrIleTyrThrLeuLysGly--MetCysProLeu 327
 Db 1753 TATTCAGTGTATAGATAAAGTTGGCCATGTATTAAGAAAAGATTAATTTGGCATTTTA 1694
 QY 328 AsnThrProGlnGlnAlaLeuAspThrProIleValValThrCysPheLeu 344
 Db 1693 TTGACTGGCAACACTGC---AGACAGCTCAAGGTATATATATTTCTC 1646
 RESULT 7
 US-08-446-648-19/c
 Sequence 19, Application US/08446648
 Patent No. 6331302
 GENERAL INFORMATION:
 APPLICANT: Genentech, Inc.
 APPLICANT: Bennett, Brian D.
 APPLICANT: Goeddel, David
 APPLICANT: Lee, James M.
 APPLICANT: Matthews, William
 APPLICANT: Tsai, Siao Ping
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.

Oy 348 ValIleLysAsnSerTyrLeuValLeu 357
 Db 2108 TTAAGCACATAACCAAGTATATTTCTT 2137

RESULT 5
 PCT-US95-05008-5
 ; Sequence 5, Application PC/TUS9505008
 ; GENERAL INFORMATION:
 ; APPLICANT: Sugen, Inc.
 ; APPLICANT: 515 Galveston Drive
 ; APPLICANT: Redwood City, California 94063-4720
 ; APPLICANT: United States of America
 ; APPLICANT: Wiscassetten E.V.
 ; APPLICANT: Hofgarten Str. 2
 ; APPLICANT: Munchen 80539
 ; APPLICANT: Germany
 ; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennle & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/05008
 ; FILING DATE: 24-APR-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/232,545
 ; FILING DATE: 22-APR-1994
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7683-074
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212)790-9090
 ; TELEFAX: (212)869-9741
 ; FAX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2770 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: DNA
 ; PCT-US95-05008-5

Alignment Scores:
 Pred. No: 4.97e-14 Length: 2770
 Score: 247.00 Matches: 99
 Percent Similarity: 44.32% Conservative: 65
 Best Local Similarity: 26.76% Mismatches: 129
 Query Match: 6.46% Indels: 78
 DB: 5 Gaps: 15

US-09-836-392-21 (1-728) x PCT-US95-05008-5 (1-2770)

Oy 15 SerGluPheArgGlnGluAlaSerMetLeuHisAlaLeuGlnHisProCysTlleValAla 34
 Db 1179 AATGACTTCCTGAGGAGGACACAGATATGAAGAACCTAGACATCCAAAGCTTATCCAG 1238

Oy 35 LeuIleGlyIleSerIle-----HisProLeuCysPheAlaLeuGlnLeuAlaProLeu 52
 Db 1239 CTTTATGCTGTTTGCACTTATGAAGATCCAAATTTATATTATTAAGAGAGTTGATGAGACAT 1298

Oy 53 SerSerLeuAsnThrValLeuSerGlnAsnAlaArgAspSerSerPheIleProLeuGly 72
 Db 1299 GGAAGTGTGCAAGATATATCTC---CAAAATGACACTGATCAAAAATC----- 1343

Oy 73 HisMetLeuThrGlnIleValle-----AlaTyrGlnIleAlaSerGlyLeuAlaTyrLeu 90
 Db 1344 ---CATCTGACTCAACAGAGTATGACATGCGGACAGGTGGCTGCTGGAATGGCCATCTG 1400

Oy 91 HisLysLysAsnIleIlePheCysAspLeuLysSerAspAsnIleLeuValTyrSerLeu 110
 Db 1401 GAGTCTCGAAGTATCAAGATCTGCTGCGCAAGAAATGTCCTC----- 1448

Oy 111 AspValLysGluHisIleAsnIleLysLeuSerAspTyrGlyIleSerArg----- 127
 Db 1449 ---GTTGGTGAACATTAATATCTACAAAGTAGAGATTTGGACTGGCCAGATTTTTAG 1505

Oy 128 -----GlnSerPheHisGlnGlyAlaLeuGlyValGlnGlyThr 140
 Db 1506 GTAGATATGAAGACATCTATGAACTGACACGAAATRAAGCTCCGCTGAAGTGACCT 1565

Oy 141 ProGlyTyrGlnAlaProGluLeuArgProArgIleValTyrAspGluLysValAspMet 160
 Db 1566 -----GCCGCCGAGGCCATTCGTAGTAATAATTCAGCATTAAGTCGATGTA 1613

Oy 161 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSer-----GlyGlnArgProAlaLeuGly 179
 Db 1614 TGGTATTGGAATCTCTTTATGAATCAATCTATCTATGCAAAATGCTTACAGTGT 1673

Oy 180 HisHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGln 199
 Db 1674 ATGACAGGTGCCACAGATATCAAGATGTGGCTCAAACTATAG-----CTTCGCGAA 1727

Oy 200 ProGluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGluCysTyrPheThrLys 219
 Db 1728 CCATCAACTGTCCA---CAGCAATTTTACAACATCAATGTTGAGTGTGATGCAAG 1784

Oy 220 ProGluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAla 239
 Db 1785 CCTAAGAGACGA-----CCTACATTTGAG 1808

Oy 240 ThrPheMetTyrGluLeuCysCysGlyLysGlnThrAlaPhePhe----- 254
 Db 1809 ACAGTGCCTTGGAACTTGAACACTTATTGAAACAGA-CCTTTATATTGAGATGCAAA 1867

Oy 255 -----SerSerGlnGlnGlnIleTyrThrValValPheTyrAspGlyLysGlu 270
 Db 1868 TAACCTCATTAAGATGACACAGCTGAGAGAAATATCAAAATTAAGTACCAAAACAAATTC 1927

Oy 271 Glu-----SerArgAsnTyrThrValValAsnThrGlnLysGlyLeuMetGluValGln 288
 Db 1928 AAATATCATCTCCAAATATCAATGTTATCAAC----- 1960

Oy 289 ArgMetCysCysProGluMetLysValSerCysGlnLeuGlnValGlnArgSerLeuTyrP 308
 Db 1961 -----CAACTGCACAATCAATCTTATCTGACA 1987

Oy 309 ThrAlaThrGluAspGlnLysIleTyrIleTyrThrLeuGlyLysGly-----MetCysProLeu 327
 Db 1988 TATCAAGATGATAGATTAAGTGGCATGATATTAAGAAAAGATTAATTGGCATTTTA 2047

Oy 328 AsnThrProGlnGlnAlaLeuAspThrProAlaValValThrCysPheLeuAlaValPro 347
 Db 2048 TTGACTGGGCAACACTGACGAGCAGCTCAAGGTATTAATGCTCACTGCTGGAAAA 2107

Oy 348 ValIleLysAsnSerTyrLeuValLeu 357
 Db 2108 TTAAGCACATAACCAAGTATATTTCTT 2137

RESULT 6
 US-08-222-616-19/c
 ; Sequence 19, Application US/08222616
 ; Patent No. 5635177

Result No.	Score	Query Match	Length	DB	ID	Description
1	856	22.4	526	4	US-09-367-212-10	Sequence 10, App
2	856	22.4	526	4	US-09-948-802-10	Sequence 10, App
3	257.5	6.7	3389	1	US-08-395-580-1	Sequence 1, Appl
4	247	6.5	2770	4	US-08-4376-509A-5	Sequence 5, Appl
5	247	6.5	2770	5	PCT-US95-05008-5	Sequence 5, Appl
6	243.5	6.4	7607	1	US-08-222-616-19	Sequence 19, App
7	243.5	6.4	7607	4	US-08-442-648-19	Sequence 19, App
8	243.5	6.4	7607	5	PCT-US95-04228-19	Sequence 19, App
9	243	6.4	1365	3	US-09-221-235-6	Sequence 6, Appl
10	243	6.4	1365	3	US-09-221-928-6	Sequence 6, Appl
11	243	6.4	1365	3	US-09-221-527-6	Sequence 6, Appl
12	243	6.4	1365	3	US-09-221-236-6	Sequence 6, Appl

13	243	6.4	1365	3	US-09-221-448-6	Sequence 6, Appl1
14	243	6.4	1365	4	US-09-221-245-6	Sequence 6, Appl1
15	243	6.4	1365	4	US-09-163-118-6	Sequence 6, Appl1
16	243	6.4	1365	4	US-09-221-528-6	Sequence 6, Appl1
17	243	6.4	1365	4	US-09-593-553-6	Sequence 6, Appl1
18	243	6.4	1365	4	US-09-221-237-6	Sequence 6, Appl1
19	243	6.4	2120	3	US-09-221-235-4	Sequence 4, Appl1
20	243	6.4	2120	3	US-09-221-928-4	Sequence 4, Appl1
21	243	6.4	2120	3	US-09-221-527-4	Sequence 4, Appl1
22	243	6.4	2120	3	US-09-221-236-4	Sequence 4, Appl1
23	243	6.4	2120	3	US-09-221-418-4	Sequence 4, Appl1
24	243	6.4	2120	4	US-09-221-245-4	Sequence 4, Appl1
25	243	6.4	2120	4	US-09-163-115-4	Sequence 4, Appl1
26	243	6.4	2120	4	US-09-221-528-4	Sequence 4, Appl1
27	243	6.4	2120	4	US-09-593-553-4	Sequence 4, Appl1
28	243	6.4	2120	4	US-09-221-237-4	Sequence 4, Appl1
29	241	6.3	3436	1	US-08-205-018-1	Sequence 1, Appl1
30	238	6.2	2359	2	US-08-469-537A-77	Sequence 77, Appl1
31	238	6.2	3358	2	US-08-466-537A-104	Sequence 104, Appl1
32	237.5	6.2	2016	4	US-09-132-118-1	Sequence 1, Appl1
33	237.5	6.2	2617	3	US-09-161-443-1	Sequence 1, Appl1
34	236	6.2	2890	5	US-07-928-464-1	Sequence 1, Appl1
35	236	6.2	2890	5	PCR-US93-07347-1	Sequence 1, Appl1
36	236	6.2	3033	1	US-08-003-311B-1	Sequence 1, Appl1
37	236	6.2	3033	1	US-08-261-432-1	Sequence 1, Appl1
38	234.5	6.1	2137	1	US-08-444-005-16	Sequence 16, Appl1
39	228.5	6.0	1788	4	US-09-529-279-14	Sequence 14, Appl1
40	228.5	6.0	2268	4	US-08-444-005-14	Sequence 14, Appl1
41	228.5	6.0	2443	2	US-08-685-625A-5	Sequence 1, Appl1
42	228.5	6.0	2656	2	US-08-685-625A-5	Sequence 5, Appl1
43	228.5	6.0	2656	4	US-09-529-279-3	Sequence 3, Appl1
44	226	5.9	4508	5	PCR-US93-06251-34	Sequence 34, Appl1
45	225	5.9	2637	4	US-09-735-934A-1	Sequence 1, Appl1

ALIGNMENTS

```

RESULT 1
US-09-387-212-10
: Sequence 10, Application US/09387212A
: Patent No 6308849
:
GENERAL INFORMATION:
APPLICANT: ROBISON, KEITH E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
: FILE REFERENCE: MNT-090
: CURRENT APPLICATION NUMBER: US/09/387,212A
: CURRENT FILING DATE: 1999-08-31
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.0
:
SEQ ID NO 10
LENGTH: 526
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
: OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-387-212-10

```

Alignment Scores:

Pred. No.:	1.e-74	Length:	536
Score:	856.00	Matches:	159
Percent Similarity:	97.69%	Conservative:	0
Best Local Similarity:	97.69%	Mismatches:	1
Query Match:	22.40%	Indels:	0
DB:	4	Gaps:	0

US-09-836-392-21 (1-728) x US-09-387-213-10 (1-526)

[illegible]

FEATURES	Location/Qualifiers
source	1. .484
	/organism="Homo sapiens"

```

US-09-836-392-21 (1-728) x BQ335590 (1-484)

QY 426 LeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeuGluProTyrMetAla 445
|||
Db 3 CTCCTTGATCATGCGACTGCTCCCTCGAGATGTGACAGGGCGTG -GAGCCCTACATGGCC 61
|||
QY 446 ProSerMetValIThrSerValValCysSerSerCugIlyArgIyGluValValTTrp 465
|||
Db 62 CCCTCCATGGTTACGTACGTGCTGTGTGACGCTCTGAGGGCAGAGGGGAGAGTGTCTGG 121
|||
QY 466 CysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThrThrTyrGlnLeuCys 485
|||
Db 122 TGCCGTGATGACAAAGGCCAACTCTTGTTGATGTACACATCCACACCTACCAAGCTGT 181
|||
QY 486 AlaArgTyrPheCysGlyValProSerProLeuArgAspMetPheProValArgProLeu 505
|||
Db 182 GCCCGGTACTCTCGCGGGGTCCCGAGCCCTCTATGACATGTTTCCCGTGGCCCTTG 241
|||
QY 506 AspHrgIuProProAlaAlaSerHisThrAlaSerProLysValProGluIlyAspSer 525
|||
Db 242 GACACGGAAACCCCGGCGACGCGACACAGGCCAACCCAAAGGTGCTGAGGGGAGCTCC 301
|||
QY 526 IleAlaAspVal-SerIleMetTyrSerGluGluLeuGluTyrGlnIleLeuIleHisG 545
|||
Db 302 ATCCGGAGACGAGGACCATCATGTACATGAGAGACTGGGACGCGAGATCTCGATCCACA 361
|||
QY 545 nglSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSerProProArgGlnAl 565
|||
Db 362 GGAAATCACTACTGACTCTCTCCATAGTCTCTTACTACTCATCCACCCCGCCAGGC 421
|||
QY 565 AlaIArgSerProSerSerSerLeuProSerSerProAlaSerSerSerValProPheSe 585
|||
Db 422 TGCCAGGTCCCCCTCAAGCCCTCCAGCTCCCAAGCAAGTCTTCACATGTGCTTCTC 481
|||
QY 585 T 585
|||
Db 482 A 482
|||

```

Search completed: April 15, 2003, 08:57:37
Job time : 1740 secs

BASE COUNT 221 a 323 c 271 g 180 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 5,64e-73 Length: 996
Score: 836.00 Matches: 172
Percent Similarity: 95.08% Conservative: 2
Best Local Similarity: 93.99% Mismatches: 5
Query Match: 21.87% Indels: 4
DB: 13 Gaps: 0

US-09-836-392-21 (1-728) x BM469840 (1-996)

QY 1 MetLeuArGHsleuArGAlaThraAspAlaMetLysAsnPheserGluPhaArgLnu 20
DB 427 ATGTGAGGACACCTCGCGGCCACCGATGCATGAAGAACTTCGCGATTCGCGAGAG 486
QY 21 AlasertMetLeuHsAlaLeuGlnHsProCysIleValAlaLeuIleGlyIleserIle 40
DB 487 GCCAGCAGCTGCACGCGCTCGACGACCCCTGCATCGTGGCGCTCATCGCATCAGCATC 546
QY 41 HisProLeuArGPhaAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
DB 547 CACCGGCTCTCTCGCGCTCGAGCTCGCGCGCTCGACGACGCTCAACACGCTGCTGTC 606
QY 61 GluAsnAlaArGAspSerSerPheIleProLeuGlyHsMetLeuThrGlnLysIleAla 80
DB 607 GAGAACGCGACAGATCTCTCTTATACCTCGGACACATGCTACCCAAATAATACCC 666
QY 81 TyrGlnIleAlaSerGlyLeuAlaIleuArGLeuHsLysLysAsnIleIlePheCysAspLeu 100
DB 667 TACCAAGATCGCTCGCGCTCGCGCTCGACGACGCGCGCTCGACGACGCGCTGCTG 726
QY 101 LysSerAspAsn-IleLeuValITPserLeuAspValLysGlnHsIleAsnIleLysLe 120
DB 727 AAGTCGACACATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
QY 120 userAspTyrGlyIleSerArGlnSerPheHsGlnGlyAlaLeuGlyValGlnGlyTh 140
DB 787 ATGTGACACGCGATTTGAGGACGATCATTCATGAGGCGCGCGCTGAGCGTGAAGGCGAC 846
QY 140 rProGlyTyrGlnAlaPro-GlnIleArGProArGlyIleValTyrAspLysValAspM 160
DB 847 TCCGTGGCTACCAAGGCG 906
QY 160 etPheSer-TyrGlyMetValLeuTyr-GluLeuLeuSerGlyGlnArGProAlaLeuG 179
DB 907 TGTTCCTCCTATGGATGTGCTCTACCGAGTTCCTGTGACGACGCGCGCGCGCGCTG 966
QY 179 y 179
DB 967 g 967

RESULT 14
AZ358773
LOCUS 707 bp DNA linear GSS 02-OCT-2000
DEFINITION M01010P17F Mouse 10kb plasmid UNGC1M library Mus musculus genomic
ACCESSION AZ358773
VERSION AZ358773.1 GI:10472473
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 707)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tinsley,A., von Niederhausern,A.
and Wright,D., Weis,R.
Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts
JOURNAL
Unpublished (2000)

COMMENT

Contact: Robert B. Weiss
University of Utah
Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0101 row: P column: 17
Seq primer: GGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 707.
Location/Qualifiers

FEATURES

source

1. 707
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UNGCM010P17"
/clone_1lb="Mouse 10kb plasmid UNGC1M library"

/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (91147321149b) (AP29072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance.

BASE COUNT 155 a 203 c 193 g 156 t
ORIGIN

Alignment Scores:

Pred. No.: 1.3e-68 Length: 707
Score: 790.00 Matches: 163
Percent Similarity: 72.29% Conservative: 17
Best Local Similarity: 65.46% Mismatches: 45
Query Match: 20.67% Indels: 24
DB: 17 Gaps: 4

US-09-836-392-21 (1-728) x AZ358773 (1-707)

QY 298 SerCysGlnLeuGlnValGlnArGSerLeuTrpAlaTrpGlnLysIleLeuYr 317
DB 5 TCGCAAGACCTTACTCTCTGAGGACACTGAG----- 37
QY 318 IleTyrThrLeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrPro 337
DB 38 -----AAGATCATGTGTGACCACTTGGACACATAGAGAG--ATGAGTGTGTTG 82
QY 338 AlaValValThrCysPheLeuAlaVal-----ProVal 348
DB 83 TCTGTGTTTCTCTCTTCTGCGATGTGACATGAAGAAATATGTGTGCTCCCTGCTTC 142
QY 349 IleLysLysAsnSerTyrLeuValLeuAlaGlyLeuAlaAspGlyLeuValAlaValPhe 368
DB 143 CCCCACAGAACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202
QY 369 ProValValArGlyThrProLysAspSerCysSerTyrLeuCysSerHsThrAlaAsn 388
DB 203 CCTGTGACG 262

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 03:48:32 ; Search time 3281 Seconds

(without alignments)
16736.115 Million cell updates/sec

Title: US-09-836-392-8_COPY_22_2205

Perfect score: 2184

Sequence: 1 atgtcagcagccctgcggc.....cttcactcgcagagaag 2184

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues 49582208

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_MA_Main:*

1: /cgn2_6/ptodata/1/pna/US0956000_COMB.seq.*

2: /cgn2_6/ptodata/1/pna/US0956001_COMB.seq.*

3: /cgn2_6/ptodata/1/pna/US0956002_COMB.seq.*

4: /cgn2_6/ptodata/1/pna/US0956003_COMB.seq.*

5: /cgn2_6/ptodata/1/pna/US0956004_COMB.seq.*

6: /cgn2_6/ptodata/1/pna/US0956005_COMB.seq.*

7: /cgn2_6/ptodata/1/pna/US0956006_COMB.seq.*

8: /cgn2_6/ptodata/1/pna/US0956007_COMB.seq.*

9: /cgn2_6/ptodata/1/pna/US0956008_COMB.seq.*

10: /cgn2_6/ptodata/1/pna/US0956009_COMB.seq.*

11: /cgn2_6/ptodata/1/pna/US0956010_COMB.seq.*

12: /cgn2_6/ptodata/1/pna/US0956011_COMB.seq.*

13: /cgn2_6/ptodata/1/pna/US0956012_COMB.seq.*

14: /cgn2_6/ptodata/1/pna/US0956013_COMB.seq.*

15: /cgn2_6/ptodata/1/pna/US0956014_COMB.seq.*

16: /cgn2_6/ptodata/1/pna/US0956015_COMB.seq.*

17: /cgn2_6/ptodata/1/pna/US0956016_COMB.seq.*

18: /cgn2_6/ptodata/1/pna/US0956017_COMB.seq.*

19: /cgn2_6/ptodata/1/pna/US0956018_COMB.seq.*

20: /cgn2_6/ptodata/1/pna/US0956019_COMB.seq.*

21: /cgn2_6/ptodata/1/pna/US0956020_COMB.seq.*

22: /cgn2_6/ptodata/1/pna/US0956021_COMB.seq.*

23: /cgn2_6/ptodata/1/pna/US0956022_COMB.seq.*

24: /cgn2_6/ptodata/1/pna/US0956023_COMB.seq.*

25: /cgn2_6/ptodata/1/pna/US0956024_COMB.seq.*

26: /cgn2_6/ptodata/1/pna/US0956025_COMB.seq.*

27: /cgn2_6/ptodata/1/pna/US0956026_COMB.seq.*

28: /cgn2_6/ptodata/1/pna/US0956027_COMB.seq.*

29: /cgn2_6/ptodata/1/pna/US0956028_COMB.seq.*

30: /cgn2_6/ptodata/1/pna/US0956029_COMB.seq.*

31: /cgn2_6/ptodata/1/pna/US0956030_COMB.seq.*

32: /cgn2_6/ptodata/1/pna/US0956031_COMB.seq.*

33: /cgn2_6/ptodata/1/pna/US0956032_COMB.seq.*

34: /cgn2_6/ptodata/1/pna/US0956033_COMB.seq.*

35: /cgn2_6/ptodata/1/pna/US0956034_COMB.seq.*

36: /cgn2_6/ptodata/1/pna/US0956035_COMB.seq.*

37: /cgn2_6/ptodata/1/pna/US0956036_COMB.seq.*

38: /cgn2_6/ptodata/1/pna/US0956037_COMB.seq.*

39: /cgn2_6/ptodata/1/pna/US0956038_COMB.seq.*

40: /cgn2_6/ptodata/1/pna/US0956039_COMB.seq.*

41: /cgn2_6/ptodata/1/pna/US0956040_COMB.seq.*

42: /cgn2_6/ptodata/1/pna/US0956041_COMB.seq.*

43: /cgn2_6/ptodata/1/pna/US0956042_COMB.seq.*

Result No.	Score	Query Match	Length	ID	Description
1	2184	100.0	3496	PCT-US00-28066-8	Sequence 8, Appl
2	2184	100.0	3496	US-09-836-392-8	Sequence 8, Appl
3	2184	100.0	3579	PCT-US01-01435-7	Sequence 37, Appl
4	2180.8	99.9	3564	PCT-US01-01435-14	Sequence 14, Appl
5	2132.4	97.6	2652	PCT-US01-08631-8042	Sequence 8042, Ap
6	1936.8	88.7	7015	US-10-132-382-5	Sequence 5, Appl
7	1936.8	88.7	7093	US-10-132-382-1	Sequence 1, Appl
8	1936.8	88.7	7209	US-10-132-382-7	Sequence 7, Appl
9	1936.8	88.7	7307	US-10-132-382-3	Sequence 3, Appl
10	1088.2	49.8	2839	PCT-US02-05109-456	Sequence 456, App
11	1088.2	49.8	2839	US-09-810-173-456	Sequence 3, App
12	941.6	43.1	995	US-09-471-275-7370	Sequence 7370, App
13	941.6	42.8	1912	US-09-637-890-0034	Sequence 9034, Ap
14	934.6	42.8	1912	US-09-644-869-9514	Sequence 9514, Ap
15	858.8	39.3	18286	PCT-US01-01354-25701	Sequence 25701, A
16	858.8	39.3	18286	US-09-764-905-25701	Sequence 38916, A
17	858.8	39.3	18286	US-09-764-905-25701	Sequence 38916, A
18	858.8	39.3	18286	US-10-092-399-25701	Sequence 25701, A
19	858.8	39.3	18286	US-10-092-399-25701	Sequence 38916, A
20	858.8	39.3	18286	PCT-US01-01354-25704	Sequence 25704, A
21	858.8	39.3	18287	US-09-836-392-8	Sequence 25704, A

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

22 858.8 39.3 18287 1 PCT-US01-01354-38919 Sequence 38919, A
23 858.8 39.3 18287 30 US-09-764-905-25704 Sequence 25704, A
24 858.8 39.3 18287 30 US-09-764-905-38919 Sequence 38919, A
25 858.8 39.3 18287 39 US-10-092-399-25704 Sequence 25704, A
26 858.8 39.3 18287 39 US-10-092-399-38919 Sequence 38919, A
27 746.6 34.2 889 19 US-09-523-654A-4 Sequence 4, Appl1
28 746.6 34.2 889 19 US-09-523-654B-4 Sequence 4, Appl1
29 746.6 34.2 889 19 US-09-902-919-4 Sequence 400, App
30 484 22.2 567 17 US-09-342-216-400 Sequence 400, App
31 484 22.2 567 17 US-09-342-216A-400 Sequence 400, App
32 484 22.2 567 34 US-09-908-601-400 Sequence 400, App
33 484 22.2 567 34 US-09-912-292-9864 Sequence 9864, Ap
34 477.6 21.9 526 17 US-09-317-311C-1275 Sequence 1275, Ap
35 477.6 21.9 526 24 US-09-637-890-101 Sequence 101, App
36 477.6 21.9 526 39 US-10-077-478-1275 Sequence 1275, Ap
37 477.6 21.9 526 40 US-10-121-925-10 Sequence 10, Appl
38 477.6 21.9 526 40 US-10-121-925-10 Sequence 10, Appl
39 409.6 18.8 567 20 US-09-539-802A-2421 Sequence 2421, Ap
40 404.8 18.5 505 18 US-09-431-517-15552 Sequence 15552, A
41 385.6 17.7 411 17 US-09-362-510A-61416 Sequence 61416, A
42 385.6 17.7 411 17 US-09-362-510B-61416 Sequence 61416, A
43 385.6 17.7 411 34 US-09-904-013-61416 Sequence 61416, A
44 382.4 17.5 397 18 US-09-489-036-27225 Sequence 27225, A
45 382.4 17.5 397 35 US-09-943-143-27225 Sequence 27225, A

ALIGNMENTS

RESULT 1
PCT-US00-28066-8
Sequence 8, Application PC/TUS0028066
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides,
TITLE OR INVENTION: Antibodies
FILE REFERENCE: F020PCT
CURRENT APPLICATION NUMBER: PCT/US00/28066
CURRENT FILING DATE: 2000-10-12
PRIOR APPLICATION NUMBER: 60/159,542
PRIOR FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: 60/165,914
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: 60/189,027
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 3496
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US00-28066-8

Query Match 100.0%; Score 2184; DB 1; Length 3496;
Best Local Similarity 100.0%; Pired, No. 0;
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAGGACCTGCGGCGCCAGCATGCGTGAAGACTTCTCGAGTTCGCGAGAG 60
Db 22 ATGCTGAGGACCTGCGGCGCCAGCATGCGTGAAGACTTCTCGAGTTCGCGAGAG 81
QY 61 GCCAGCATGTCGACGCGCTGACGACCCCTGATCGTGGGCTCATGGCATCAGCATC 120
Db 82 GCCAGCATGTCGACGCGCTGACGACCCCTGATCGTGGGCTCATGGCATCAGCATC 141
QY 121 CACCGCTCTGCTTGGCTGAGTGCAGCGCGCTCAGACAGCTTCACACCGTGTCTCC 180
Db 142 CACCGCTCTGCTTGGCTGAGTGCAGCGCGCTCAGACAGCTTCACACCGTGTCTCC 201
QY 181 GAGAACGCGAGAGATTCTTCTTATACCCCTGGGACACATGCTCACCCAAAATAATACC 240
Db 202 GAGAACGCGAGAGATTCTTCTTATACCCCTGGGACACATGCTCACCCAAAATAATACC 261

QY 241 TACAGATCGCTGCGGCGCTTACCTGACACAGAAAAATCATCTTCTGTGACTG 300
Db 262 TACAGATCGCTGCGGCGCTTACCTGACACAGAAAAATCATCTTCTGTGACTG 321
QY 301 AAGTCGACAAATCTTGTGTGTGCTCCCTTACGTCAAGAGACATCATCACTAAGCTA 360
Db 322 AAGTCGACAAATCTTGTGTGTGCTCCCTTACGTCAAGAGACATCATCACTAAGCTA 381
QY 361 TCTGACTACGGGATTTGAGGACATCATTCATGAGGGCGCCATGAGGTGAGGACT 420
Db 382 TCTGACTACGGGATTTGAGGACATCATTCATGAGGGCGCCATGAGGTGAGGACT 441
QY 421 CCGTGTACACAGGCGCCAGAGATCAGGCTTCATTTGATATGATGAGAGATGATG 480
Db 442 CCGTGTACACAGGCGCCAGAGATCAGGCTTCATTTGATATGATGAGAGATGATG 501
QY 481 TTCCTCTATGAAATGTGCTCTACAGTGTGCTGACAGACGCGCTGACCTGGCCAC 540
Db 502 TTCCTCTATGAAATGTGCTCTACAGTGTGCTGACAGACGCGCTGACCTGGCCAC 561
QY 541 CACAGGCTCCAGATTTGCGCAAGAGCTGCAAGGGCANTCCGCGTTCTGGGCGACCG 600
Db 562 CACAGGCTCCAGATTTGCGCAAGAGCTGCAAGGGCANTCCGCGTTCTGGGCGACCG 621
QY 601 GAGAGATGACAGTTCGCGCGACTGACAGCGCTCATGTAGATGAGTCTGGAGACTAAGCCA 660
Db 622 GAGAGATGACAGTTCGCGCGACTGACAGCGCTCATGTAGATGAGTCTGGAGACTAAGCCA 681
QY 661 GAGAGAGGACCGCTGCGCTGCTGCTGCTGAGCCAGATGAAGACCCGACTTTTCCACC 720
Db 682 GAGAGAGGACCGCTGCGCTGCTGCTGCTGAGCCAGATGAAGACCCGACTTTTCCACC 741
QY 721 TTCATGATGAACGTGCTGTGAGAGAGAGACGCTTCTCATCCAGGCGCAGAG 780
Db 742 TTCATGATGAACGTGCTGTGAGAGAGAGACGCTTCTCATCCAGGCGCAGAG 801
QY 781 TACACCGTGTGTTTGGATGGAAGAGAGATGCAAGAACTACAGGCTGTGAACACA 840
Db 802 TACACCGTGTGTTTGGATGGAAGAGAGATGCAAGAACTACAGGCTGTGTGAACACA 861
QY 841 GAGAGAGGCTCATGAGATGACAGATGATGCTGCTGAGGATGAAGTGTGCTGAC 900
Db 862 GAGAGAGGCTCATGAGATGACAGATGATGCTGCTGAGGATGAAGTGTGCTGAC 921
QY 901 CTCAGGCTCCAGATCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 922 CTCAGGCTCCAGATCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981
QY 961 CTCAGGCTCCAGATCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 982 CTCAGGCTCCAGATCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1041
QY 1021 ACCTGCTTCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1042 ACCTGCTTCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
QY 1081 GCCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 1102 GCCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161
QY 1141 TACCTGTCTACACAG 1200
Db 1162 TACCTGTCTACACAG 1221
QY 1201 CAGAACCCCTTACAG 1260
Db 1222 CAGAACCCCTTACAG 1281
QY 1261 AGCAATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db 1282 AGCAATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1341
QY 1321 GAGCCTACATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380

[illegible]

```

RESULT 2
US-09-836-392-8
; Sequence 8, Application US/09836392
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides, &
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PTO20P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
;

```

```

; PRIOR FILING DATE: 1999-10-15.
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17.
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-392-8

Query Match      100.0%; Score 2184; DB 32; Length 3496;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	1	ATGCGAGGACATCTCGGGGGCCAGGAGCCATGAAGAACTTCTCCAGATGTTCCGGCAGAG	60
Db	22	ATGCGAGGACATCTCGGGGGCCAGGAGTGCATGAAGAACTTCTCCAGATGTTCCGGCAGAG	81
QY	61	GCCAGCATGCTGCACAGGGCTGCAGACACCCCTGCATGCTGGCGCTCATCGGCATCAGATC	120
Db	82	GCCAGCATGCTGCACAGGGCTGCAGACACCCCTGCATGCTGGCGCTCATCGGCATCAGATC	141
QY	121	CACCGCGCTGCTGGCCCTGGAGACTGGGGCGGCTGAGAGGCTCAACAACCGTGTCTC	180
Db	142	CACCGCGCTGCTGGCCCTGGAGACTGGGGCGGCTGAGAGGCTCAACAACCGTGTCTC	201
QY	181	GAGAACGCCAGAGATTTCTCTTTATACCCCTGGGACACATGCTCAACCCAAAATAATAGC	240
Db	202	GAGAACGCCAGAGATTTCTCTTTATACCCCTGGGACACATGCTCAACCCAAAATAATAGC	261
QY	241	TACCGATGCGCTGGGGCTGGCTTACCTGACACAAGAAAAACATCATCTTCTGTGACCTG	300
Db	262	TACCGATGCGCTGGGGCTGGCTTACCTGACACAAGAAAAACATCATCTTCTGTGACCTG	321
QY	301	AAGTGGGACACATCTGCTGTGGTGCCCTTACGCTCAAGAGAGCATCAATCAATCAAGCTA	360
Db	322	AAGTGGGACACATCTGCTGTGGTGCCCTTACGCTCAAGAGAGCATCAATCAATCAAGCTA	381
QY	361	TCTGACTACGGGATTTGAGGCGACGTCATTTCCATGAGGGCGCCCTAGGCGTGGAGGCACT	420
Db	382	TCTGACTACGGGATTTGAGGCGACGTCATTTCCATGAGGGCGCCCTAGGCGTGGAGGCACT	441
QY	421	CCTGGCTACAGAGCCCCAGAGATCAGGGCTTGGCATTTATATATATAGAAAGTATAGATG	480
Db	442	CCTGGCTACAGAGCCCCAGAGATCAGGGCTTGGCATTTATATATATAGAAAGTATAGATG	501
QY	481	TTCCTCTATGGAATGCTGCTACAGATTTGCTGCAGAGACAGCCCTGCACTAGGGCCAC	540
Db	502	TTCCTCTATGGAATGCTGCTACAGATTTGCTGCAGAGACAGCCCTGCACTAGGGCCAC	561
QY	541	CACCAAGCTCAGATTTGCCAAGAAAGCTGTCCAAGGGCAATCGCCCGGTTCTGGGGCAGCG	600
Db	562	CACCAAGCTCAGATTTGCCAAGAAAGCTGTCCAAGGGCAATCGCCCGGTTCTGGGGCAGCG	621
QY	601	GAGGAGTGCATTTCCGGGAGACTGCAGAGCGCTGCATGATGATGAGTGCCTGGGACACTAAGCCA	660
Db	622	GAGGAGTGCATTTCCGGGAGACTGCAGAGCGCTGCATGATGATGAGTGCCTGGGACACTAAGCCA	681
QY	661	GAGAAAGCAGCCGCTGGCCCTGTCCGCTGTGAGCCAGATGAAGAGCCGACATTTTGGCCACC	720
Db	682	GAGAAAGCAGCCGCTGGCCCTGTCCGCTGTGAGCCAGATGAAGAGCCGACATTTTGGCCACC	741
QY	721	TTCATGTATGAACCTGTGCTGTGGGAGCAGACAGCTTCTTCTCATCCCAAGGCCAGGAG	780
Db	742	TTCATGTATGAACCTGTGCTGTGGGAGCAGACAGCTTCTTCTCATCCCAAGGCCAGGAG	801
QY	781	TACCCGTGTGTTTTGGGATGGAAGAAAGAGAGTCCAGGAATCTACAGGTGTGTAACACA	840
Db	802	TACCCGTGTGTTTTGGGATGGAAGAAAGAGAGTCCAGGAATCTACAGGTGTGTAACACA	861

OY	841	GAGAAAGGGCCCATGAGAGGGGACAGAGGATGTGCTGCCCGGGATGGAAGGTGAGCTGGACG	900
Db	862	GAGAAAGGGCCTCATGAGGTGACAGAGGATGTGCTGCCCGGGATGGAAGGTGAGCTGGACG	921
OY	901	CTCCAGGTCAGAGATATCCCTGTGTGGACAGGCCACCGAGAGACAGAAAAATCTACATCTACACC	960
Db	922	CTCCAGGTCAGAGATATCCCTGTGTGGACAGGCCACCGAGAGACAGAAAAATCTACATCTACACC	981
OY	961	CTCAAGGGCATGTGCCCCCTTAAACACACCCCAACAGGGCTTGGATATCTCAGCTGTGCTC	1020
Db	982	CTCAAGGGCATGTGCCCCCTTAAACACACCCCAACAGGGCTTGGATATCTCAGCTGTGCTC	1041
OY	1021	ACCGCTCTTGGGCGGCGCTGTATTAATAAANAATCTTACTCTGTGCTTACGGGGCTC	1080
Db	1042	ACCGCTCTTGGGCGGCGCTGTATTAATAAANAATCTTACTCTGTGCTTACGGGGCTC	1101
OY	1081	GCCCATGGGCTTGTGGGCTGTGTTCCTCGTGTGTCGGGGACACCCAAAAGAGACGTGTCTC	1140
Db	1102	GCCCATGGGCTTGTGGGCTGTGTTCCTCGTGTGTCGGGGACACCCAAAAGAGACGTGTCTC	1161
OY	1141	TACCTGTGCTACACACACAGCCAAAGGTCCAAAGTTACAGATCGCGGATGAAAGACACAGG	1200
Db	1162	TACCTGTGCTACACACAGCCAAAGGTCCAAAGTTACAGATCGCGGATGAAAGACACAGG	1221
OY	1201	CAGAAACCCCTAACCCAGTAAGAGGGCCATGAGAGTGTAAACAGGGGCTCTGAGGTCTGTATC	1260
Db	1222	CAGAAACCCCTAACCCAGTAAGAGGGCCATGAGAGTGTAAACAGGGGCTCTGAGGTCTGTATC	1281
OY	1261	AGCAATGGGCGCGGCGCTCTCTTGTATCATGATGTGCTCCCTCGAAGATCTGACAGGGGCTG	1320
Db	1282	AGCAATGGGCGCGGCGCTCTCTTGTATCATGATGTGCTCCCTCGAAGATCTGACAGGGGCTG	1341
OY	1321	GAGCCCTACATGAGCCCCCTCCATGTTACGTCAGTGCTGTGACGCTCTGAGGGCAGAGGG	1380
Db	1342	GAGCCCTACATGAGCCCCCTCCATGTTACGTCAGTGCTGTGACGCTCTGAGGGCAGAGGG	1401
OY	1381	GAGAGAGTCCGTGTGTGCTCCCTGATGACAGAGGCCAACTCTTGTGTATGTACATCTCCACC	1440
Db	1402	GAGAGAGTCCGTGTGTGCTCCCTGATGACAGAGGCCAACTCTTGTGTATGTACATCTCCACC	1461
OY	1441	ACCTACACAGTGTGTGCGCGGCTACTCTGTGGGGGTCCCCAGCCCCCTCAGGGACATGTTT	1500
Db	1462	ACCTACACAGTGTGTGCGCGGCTACTCTGTGGGGGTCCCCAGCCCCCTCAGGGACATGTTT	1521
OY	1501	CCCGTGGGGCCCTTGGACACAGGAACCCCGCGGACGACAGCCACAGGGCCAAACCCAAAGGTG	1560
Db	1522	CCCGTGGGGCCCTTGGACACAGGAACCCCGCGGACGACAGCCACAGGGCCAAACCCAAAGGTG	1581
OY	1561	CCTAGAGGGGACCTCCATCCGGGGGACGTGTAGATATATTAACATGTAGAGAGCTGGGACACGAG	1620
Db	1582	CCTAGAGGGGACCTCCATCCGGGGGACGTGTAGATATATTAACATGTAGAGAGCTGGGACACGAG	1641
OY	1621	ATTCGTATTCACACAGAAATCACTACATGACTACTGTGTCATGTCGCTACACCTACATCC	1680
Db	1642	ATTCGTATTCACACAGAAATCACTACATGACTACTGTGTCATGTCGCTACCTCTCATCC	1701
OY	1681	CCACCCCGGACAGGCTGCGAGGTCCCCCTCAAGCCTCCACAGCTCCCAACAGATCTTCC	1740
Db	1702	CCACCCCGGACAGGCTGCGAGGTCCCCCTCAAGCCTCCCAACAGATCTTCC	1761
OY	1741	AGTGTGCTTCTTCCACACCGACTGCGAGAGACTAGACATGCTACATACGCGCGGTCTGTCC	1800
Db	1762	AGTGTGCTTCTTCCACACCGACTGCGAGAGACTAGACATGCTACATACGCGCGGTCTGTCC	1821
OY	1801	TCCGACAGGTGTGACATGTACCTGACCCCCATGGAGGGGAGACCTTCAGCCAGACAGCTG	1860
Db	1822	TCCGACAGGTGTGACATGTACCTGACCCCCATGGAGGGGAGACCTTCAGCCAGACAGCTG	1881
OY	1861	CAGGCGGTGAAGATCTCGCCGCTCAGAGACTCTATTGTGGTCCCAAGGCGCGGTGGAGAT	1920
Db	1882	CAGGCGGTGAAGATCTCGCCGCTCAGAGACTCTATTGTGGTCCCAAGGCGCGGTGGAGAT	1941
OY	1921	GTATGTGCTATGTGCTGGAGAAAGATCTTGAAGCCACAGGGGGCGAGCTATTGCGCTC	1980

Accession	Sequence	Position
Db 1942	GTATTGCTCATTTGGCCCTGGAGAAAGATTCTGAAGCCACGGGGCCGAATCAATTCGGCTC	2001
Qy 1981	TTAAAGCCCGAGAGACTGACTCCGCATGGGGTCTGTGTGATGTCTCCGTGTGTGCAAG	2040
Db 2002	TTAAAGCCCGAGAGCTGACTCCGCATGGGGTCTGTGTGATGTCTCCGTGTGTGCAAG	2061
Qy 2041	GACACTGTGTGTGACACTTTGAAATGAAACACAGATGTGTGCTGTGCTCTGTGAGG	2100
Db 2062	GACACTGTGTGTGACACTTTGAAATGAAACACAGATGTGTGCTGTGCTCTGTGAGG	2121
Qy 2101	GGCTGGGAGCGCAGGGAGTTCGACATTTTCTACAGTCTTACGAGAGAGCTGGGCGCGGTG	2160
Db 2122	GGCTGGGAGCGCAGGGAGTTCGACATTTTCTACAGTCTTACGAGAGAGCTGGGCGCGGTG	2181
Qy 2161	GAGGCTTGCACTGCGCAAGAAAGG	2184
Db 2182	GAGGCTTGCACTGCGCAAGAAAGG	2205

```

RESULT 3
PCT-US01-01435-37
: Sequence 37, Application PC/TUS0101435
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: 21 human secreted proteins
: FILE REFERENCE: P5725PCT
: CURRENT APPLICATION NUMBER: PCT/US01/01435
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION NUMBER: 60/226,282
: PRIOR FILING DATE: 2000-08-18
: NUMBER OF SEQ ID NOS: 145
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 37
: LENGTH: 3579
: TYPE: DNA
: ORGANISM: Homo sapiens
PCT-US01-01435-37

```

Query Match	100.0%	Score 2184;	DB 1;	Length 3579;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2184; Conservative	0;	Mismatches	0;	Gaps 0;

OY	1	ATGCGTAGGACATCGCGGGGACCGGATGCGCATGAAAGAACTTCTCGGAGTTCCGGCAGAG	60
Db	68	ATGCTAGGACATCGCGGGACCGGATGCGCATGAAAGAACTTCTCGGAGTTCCGGCAGAG	120
OY	61	GCGAGCATGCTGCAGCGGCTCAGACACCCCGCATCGGGGCTCATCGGATGAGCATC	120
Db	128	GCGAGCATGCTGCAGCGGCTCAGACACCCCGCATCGGGGCTCATCGGATGAGCATC	180
OY	121	CACCGGCTCTGCTTGGCCCTGGAGCTGGCGGCTCAGAGAGCTCAACACCGTCTGCC	180
Db	188	CACCGGCTCTGCTTGGCCCTGGAGCTGGCGGCTCAGAGAGCTCAACACCGTCTGCC	240
OY	181	GAGAACGGCAGATGTTCTTATATACCTGGGACATGCTCACCCAAAATATGCC	240
Db	248	GAGAACGGCAGATGTTCTTATATACCTGGGACATGCTCACCCAAAATATGCC	300
OY	241	TACAGATCGCTGGGGCTGGCTTACCTGGCAAGAAAAATCATCTCTGGAGCTG	300
Db	308	TACAGATCGCTGGGGCTGGCTTACCTGGCAAGAAAAATCATCTCTGGAGCTG	360
OY	301	AAGTGGACAACATTCTGTGTGTGCTCCCTTACGCTCAAGGAGCAATCAACATTAAGCTA	360
Db	368	AAGTGGACAACATTCTGTGTGTGCTCCCTTACGCTCAAGGAGCAATCAACATTAAGCTA	420
OY	361	TCTGACTACGGGATTTGAGGCACTCATTTCCATGAGGGCGCCCTAGGGCTGAGAGGCACT	420
Db	428	TCTGACTACGGGATTTGAGGCACTCATTTCCATGAGGGCGCCCTAGGGCTGAGAGGCACT	480
OY	421	CTTGCTACACAGGCCCAAGAGATCAGGCTCTGCATTGTATGTATGAAAGTAGATATG	480

D	b	488	CTGGCTACCAAGGCCCCCAAGATCAAGGCGCTCCGACTTGTATATAGTAGAAGTAGATGATG	547
O	y	481	TTCTCCTATGGAATGGTGGCTCTACAGATGGTCTGTACAGACAGCGCCCTGCACCTGGCCAC	540
D	b	548	TTCTCCTATGGAATGGTGGCTCTACAGATGGTCTGTGTACAGACAGCGCCCTGCACCTGGCCAC	607
O	y	541	CACAGCTCCAGATTGCGCAAGAAAGCTGTCCAGAGGGCATCCGGCCGGTCTGTGGGACAGCG	600
D	b	608	CACAGCTCCAGATTGCGCAAGAAAGCTGTCCAGAGGGCATCCGGCCGGTCTGTGGGACAGCG	667
O	y	601	GAGGAATGCAATGCCGGGAGACTCAGGCGCTCATGATGAGTGGTGGAGCACTAAAGCA	660
D	b	668	GAGGAATGCAATGCCGGGAGACTCAGGCGCTCATGATGAGTGGTGGAGCACTAAAGCA	727
O	y	661	GAGAGCGACCGCTGGCCCTGTGGTGTGTGAGCCAGATGAAAGCACCGACATTTTGCACC	720
D	b	728	GAGAGCGACCGCTGGCCCTGTGGTGTGTGAGCCAGATGAAAGCACCGACATTTTGCACC	787
O	y	721	TTCTATGATGAATGAGTGGCTGTGGGAAGCAGACAGCGCTTCTCATCTCCAGGGCCAGAG	780
D	b	788	TTCTATGATGAATGAGTGGCTGTGGGAAGCAGACAGCGCTTCTCATCTCCAGGGCCAGAG	847
O	y	781	TACACCTGGTGTGTTTGGATGGAAGAAAGAGAGATCCAGAACTACACGGTGGTGAACAA	840
D	b	848	TACACCTGGTGTGTTTGGATGGAAGAAAGAGAGATCCAGAACTACACGGTGGTGAACAA	907
O	y	841	GAGAGGGCCCTCATGAGATGTCAGAGAGATGTGCTGCCCTGGGATGAAGTGAAGTCCAG	900
D	b	908	GAGAGGGCCCTCATGAGATGTCAGAGAGATGTGCTGCCCTGGGATGAAGTGAAGTCCAG	967
O	y	901	CTCCAGTCCAGAGATCCCGTGTGGAGACAGCCACCCAGAGACCAAAATATACATCTACACC	960
D	b	968	CTCCAGTCCAGAGATCCCGTGTGGAGACAGCCACCCAGAGACCAAAATATACATCTACACC	1027
O	y	961	CTCAAGGGCATGTGCCCCCTTAAACACACCCCAACAGGCTTGATATCTCAGCTGTGCTC	1020
D	b	1028	CTCAAGGGCATGTGCCCCCTTAAACACACCCCAACAGGCTTGATATCTCAGCTGTGCTC	1087
O	y	1021	ACGTGCTTCTGGCGCTGCCCTGTTATTTAAAGAAATTCCTACCTGATCTTACGGGGCTC	1080
D	b	1088	ACGTGCTTCTGGCGCTGCCCTGTTATTTAAAGAAATTCCTACCTGATCTTACGGGGCTC	1147
O	y	1081	GCCCATGGGCTTGTGGCTGTGTTTCCCGTGTGGCGGGCACCCCAAGAGCAGCTGTCTCC	1140
D	b	1148	GCCCATGGGCTTGTGGCTGTGTTTCCCGTGTGGCGGGCACCCCAAGAGCAGCTGTCTCC	1207
O	y	1141	TACGTGTGTACACACAGGCAACAGCAACAGTCCAAAGTTAGATGCGGATGAAGACGACAG	1200
D	b	1208	TACGTGTGTACACACAGGCAACAGCAACAGTCCAAAGTTAGATGCGGATGAAGACGACAG	1267
O	y	1201	CAGAAACCCCTACCCAGTGAAGAGCCATGAGAGTGTCTAAACAGGCGCTGTGAGTGTGTATC	1260
D	b	1268	CAGAAACCCCTACCCAGTGAAGAGCCATGAGAGTGTCTAAACAGGCGCTGTGAGTGTGTATC	1327
O	y	1261	AGCAATGGGCGGGGCTCCTTGTATCGACTGTGCCCTCCGTGGAATGTGACGCGGGCTG	1320
D	b	1328	AGCAATGGGCGGGGCTCCTTGTATCGACTGTGCCCTCCGTGGAATGTGACGCGGGCTG	1387
O	y	1321	GAGCGCCATCATATGGCCCGCCCTCCATAGTGTAGTCACTGTGTGTGACGCTCTGAAGGACAGAG	1380
D	b	1388	GAGCGCCATCATATGGCCCGCCCTCCATAGTGTAGTCACTGTGTGTGACGCTCTGAAGGACAGAG	1447
O	y	1381	GAGAGAGTGTGATGTGCTGTGCTGTGATGACAAGAGCCAACTCTTGATGTATACCATCCAC	1440
D	b	1448	GAGAGAGTGTGATGTGCTGTGCTGTGATGACAAGAGCCAACTCTTGATGTATACCATCCAC	1507
O	y	1441	ACCTACCAAGCTGTGTGCGCGGTACTTCTGTGGGGGTCCCAAGCCCGCTCTAGGGACATGTT	1500
D	b	1508	ACCTACCAAGCTGTGTGCGCGGTACTTCTGTGGGGGTCCCAAGCCCGCTCTAGGGACATGTT	1567
O	y	1501	CCCGTGGGGCCCTTGTGACACAGGAACCCCGGACACCAAGCCCAACCCAAAGGTG	1566
D	b	1568	CCCGTGGGGCCCTTGTGACACAGGAACCCCGGACACCAAGCCCAACCCAAAGGTG	1627

OY	1561	CGTGAAGGGGAGCTCATTCCGGGACGCGAGCATCATCTAAGAGAGACTGGGACACGAG	1620
Dp	1628	CTTGAAGGGGAGCTCATTCCGGGACGCGAGCATCATCTAAGAGAGACTGGGACACGAG	1687
OY	1621	ATTCGTATCCACGAGAAATCACTCACTGACTACTGCTCCATGCTCCCTACTCCTCATCC	1680
Dp	1688	ATTCGTATCCACGAGAAATCACTCACTGACTACTGCTCCATGCTCCCTACTCCTCATCC	1747
OY	1681	CCACCCCGCCAGAGCTGGCAGAGTCCCCCTCAAGCTTCCCACTCCCCACGAAAGTCTCC	1740
Dp	1748	CCACCCCGCCAGAGCTGGCAGAGTCCCCCTCAAGCTTCCCACTCCCCACGAAAGTCTTC	1807
OY	1741	AGTGTGCTTTCTCCACACGAGCTCGAGAGACTCAGACATGCTACATACGCCGCTGCTGC	1800
Dp	1808	AGTGTGCTTTCTCCACACGAGCTCGAGAGACTCAGACATGCTACATACGCCGCTGCTGC	1867
OY	1801	TCCGACAGGCTGTGAGCATGACTGACCTGACCCCCATGAGCGGGAGACCTTCACGCCAGCCTG	1860
Dp	1868	TCCGACAGGCTGTGAGCATGACTGACCTGACCCCCATGAGCGGGAGACCTTCACGCCAGCCTG	1927
OY	1861	CAGGCGCTGAAAGATCCCTGCGCGCTCAGAGACCTATTGGGTGCCACAGGGGGGTGGAAAT	1920
Dp	1928	CAGGCGCTGAAAGATCCCTGCGCGCTCAGAGACCTATTGGGTGCCACAGGGGGGTGGAAAT	1987
OY	1921	GTTATGCTATTGGGCTCGAGAGAGATTCTGAAGCCACGCGGGGCGGACTATTGCGCTC	1980
Dp	1988	GTTATGCTATTGGGCTCGAGAGAGATTCTGAAGCCACGCGGGGCGGACTATTGCGCTC	2047
OY	1981	TTTAAAGCCCGAGAGCTGACTCCGCGATGGGGTCTGCTGGATGCTGCTGGTGGCAAG	2040
Dp	2048	TTTAAAGCCCGAGAGCTGACTCCGCGATGGGGTCTGCTGGATGCTGCTGGTGGCAAG	2107
OY	2041	GACACTGTGTGCTCACCTTTGAAAAATGAAAACACAGAGAGTGCTGCGCTCTGAGAG	2100
Dp	2108	GACACTGTGTGCTCACCTTTGAAAAATGAAAACACAGAGAGTGCTGCGCTCTGAGAG	2167
OY	2101	GGCTGGGGCGCCAGAGAGTTTCGACATTTTCTACAGTCTCTACGAGAGAGCTGGGCGGCTG	2160
Dp	2168	GGCTGGGGCGCCAGAGAGTTTCGACATTTTCTACAGTCTCTACGAGAGAGCTGGGCGGCTG	2227
OY	2161	GAGGCTTGCACTGCGCAGAGAGG	2184
Dp	2228	GAGGCTTGCACTGCGCAGAGAGG	2251

```

RESULT 4
PCT-US01-01435-14
Sequence 14, Application PC/TUS0101435
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: 21 human secreted proteins
FILE REFERENCE: P5726PCT
CURRENT APPLICATION NUMBER: PCT/US01/01435
CURRENT FILING DATE: 2001-01-17
PRIOR FILING DATE: 2000-08-18
NUMBER OF SEQ ID NOS: 145
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 3564
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US01-01435-14

```

Query Match	99.9%	Score 2180.8	DB 1	Length 3564
Best Local Similarity	99.9%	Pred No. 0		
Matches 2182	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Q7	1	ATGCTGAGGCACTTGGCGGCCACCGATGCGATGAGAACTTCTCCGAGTTCGGCGCAGAG	60	
Db	74	ATGCTGAGGCACTTGGCGGCCACCGATGCGATGAGAACTTCTCCGAGTTCGGCGCAGAG	133	

QY 61 GCCAGATGCTGACGCGCTGCAGACCCCTGATCGTGGGCTCATCGCATGACATC 120
| | | | |
Db 134 GCCAGATGCTGACGCGCTGCAGACCCCTGATCGTGGGCTCATCGCATGACATC 193
QY 121 CACCGCTCTCTGCGCTGAGCTGCGCGCTCAGACCTTCACACCGTGTCTCC 180
| | | | |
Db 194 CACCGCTCTCTGCGCTGAGCTGCGCGCTCAGACCTTCACACCGTGTCTCC 253
QY 181 GAGAACGCCAGAGATTCCTCTTATACCCCTGGAGACATGCTACCCAAAATATACC 240
| | | | |
Db 254 GAGAACGCCAGAGATTCCTCTTATACCCCTGGAGACATGCTACCCAAAATATACC 313
QY 241 TACAGATGCGCTGCGGCTGACCTACCTGCACAAAGAAACATCATCTTGTGACCTG 300
| | | | |
Db 314 TACAGATGCGCTGCGGCTGACCTACCTGCACAAAGAAACATCATCTTGTGACCTG 373
QY 301 AAGTGGCAACATTCCTGCTGCTGCTTACCTGCAAGAGACATCAATCAAGCTA 360
| | | | |
Db 374 AAGTGGCAACATTCCTGCTGCTGCTTACCTGCAAGAGACATCAATCAAGCTA 433
QY 361 TCTGACTACGGGATTTGAGGCAATTCATGAGGGGCGCTAGGCGTGAAGGCGACT 420
| | | | |
Db 434 TCTGACTACGGGATTTGAGGCAATTCATGAGGGGCGCTAGGCGTGAAGGCGACT 493
QY 421 CCTGGCTACAGGCGCCAGAGATCAGGCTGCAATTTATATGATGAGAGATGATATG 480
| | | | |
Db 494 CCTGGCTACAGGCGCCAGAGATCAGGCTGCAATTTATGATGAGAGATGATATG 553
QY 481 TTTCTCTATGGAATGCTCTACAGATTCCTGCTGAGAGACGCGCTGCACCTGGGCGAC 540
| | | | |
Db 554 TTTCTCTATGGAATGCTCTACAGATTCCTGCTGAGAGACGCGCTGCACCTGGGCGAC 613
QY 541 CACAGCTCAGATTCGCAAGAGCTGTCAGAGGCTCCCGGCTCTGGGGAGGCG 600
| | | | |
Db 614 CACAGCTCAGATTCGCAAGAGCTGTCAGAGGCTCCCGGCTCTGGGGAGGCG 673
QY 601 GAGAAAGCTGATCCGCGCTGAGCTGAGCTCATGATGAGTGTGGAGACTAAGCA 660
| | | | |
Db 674 GAGAAAGCTGATCCGCGCTGAGCTGAGCTCATGATGAGTGTGGAGACTAAGCA 733
QY 661 GAGAAAGCTGATCCGCGCTGAGCTGAGCTCATGATGAGTGTGGAGACTAAGCA 720
| | | | |
Db 734 GAGAAAGCTGATCCGCGCTGAGCTGAGCTCATGATGAGTGTGGAGACTAAGCA 793
QY 721 TTTATGATGAACTGCTGCTGGGAGAGACAGGCTCTTCTCCAGGCGCCAGAG 780
| | | | |
Db 794 TTTATGATGAACTGCTGCTGGGAGAGACAGGCTCTTCTCCAGGCGCCAGAG 853
QY 781 TACACCTGCTGCTTTTGGGATGGAAGAGAGTCCAGAACTACACGCTGTGAACA 840
| | | | |
Db 854 TACACCTGCTGCTTTTGGGATGGAAGAGAGTCCAGAACTACACGCTGTGAACA 913
QY 841 GAGAAAGCTGATGAGCTGAGAGATGCTGCTCCCTGGGATGAGTGTGAGCTG 900
| | | | |
Db 914 GAGAAAGCTGATGAGCTGAGAGATGCTGCTCCCTGGGATGAGTGTGAGCTG 973
QY 901 CTCAGGCTCAGAGATTCCTGCTGAGAGACACGAGGACCAAGAAATCTACATCTACACC 960
| | | | |
Db 974 CTCAGGCTCAGAGATTCCTGCTGAGAGACACGAGGACCAAGAAATCTACATCTACACC 1033
QY 961 CTCAGGCTCAGAGATTCCTGCTGAGAGACACGAGGCTTGGATCTACACGCTGCTC 1020
| | | | |
Db 1034 CTCAGGCTCAGAGATTCCTGCTGAGAGACACGAGGCTTGGATCTACACGCTGCTC 1093
QY 1021 ACCTGCTCTTGGCGCTGCTGCTTATTAAGAAATCTTACCTGTCTTACGGGCTC 1080
| | | | |
Db 1094 ACCTGCTCTTGGCGCTGCTGCTTATTAAGAAATCTTACCTGTCTTACGGGCTC 1153
QY 1081 GCCGATGGGCTGTGGCTGTGTTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
| | | | |
Db 1154 GCCGATGGGCTGTGGCTGTGTTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1213
QY 1141 TACCTGTCTCACAACAGCCAAAGGTCCAAGTTCCAGATCGCGATGAAAGCGACGG 1200

Db 1214 TACCTGTCTCACAACAGCCAAAGGTCCAAGTTCCAGATCGCGGATGAAAGAGCGACGG 1273
QY 1201 CAGAACCCCTACCCAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
| | | | |
Db 1274 CAGAACCCCTACCCAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1333
QY 1261 AGCAATGGGCGGCGCTCTGCTGATGAGTGTGCTCTCTCTGAGAGTCTGAGGCGGCTG 1320
| | | | |
Db 1334 AGCAATGGGCGGCGCTCTGCTGATGAGTGTGCTCTCTCTGAGAGTCTGAGGCGGCTG 1393
QY 1321 GAGCCCTACATGCGCCCTCCATGCTTACGTCAGTGTGCTGACCTGTGAGGCGAGAGG 1380
| | | | |
Db 1394 GAGCCCTACATGCGCCCTCCATGCTTACGTCAGTGTGCTGACCTGTGAGGCGAGAGG 1453
QY 1381 GAGAGAGTCTGCTGCTGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
| | | | |
Db 1454 GAGAGAGTCTGCTGCTGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1513
QY 1441 ACCTACAGCTGTGCTGCTGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
| | | | |
Db 1514 ACCTACAGCTGTGCTGCTGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1573
QY 1501 CCGGTGGGCGCTTGGAGACGGAACCCCGGAGCGACGACGAGGCGCAACCAAGGAGT 1560
| | | | |
Db 1574 CCGGTGGGCGCTTGGAGACGGAACCCCGGAGCGACGACGAGGCGCAACCAAGGAGT 1633
QY 1561 CCTAGAGGAGGAGTCTGCTGCTGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTG 1620
| | | | |
Db 1634 CCTAGAGGAGGAGTCTGCTGCTGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTG 1693
QY 1621 ATCTGATCCACAGGAATCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
| | | | |
Db 1694 ATCTGATCCACAGGAATCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1753
QY 1681 CCACCCCGCAGGCTGCCAGGTCCCGCTCAAGCTCCCGACCTCCCGACGAGAGTCTTCC 1740
| | | | |
Db 1754 CCACCCCGCAGGCTGCCAGGTCCCGCTCAAGCTCCCGACGAGAGTCTTCC 1813
QY 1741 AGTGTGCTTCTCTCCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1800
| | | | |
Db 1814 AGTGTGCTTCTCTCCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1873
QY 1801 TCCGACAGGCTGAGACGATGCTGACCCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
| | | | |
Db 1874 TCCGACAGGCTGAGACGATGCTGACCCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1933
QY 1861 CAGGCGGTGAAGATCTGCGCTGAGAGCTCATTTTGGTCCCGAGGCGGCTGAGAT 1920
| | | | |
Db 1934 CAGGCGGTGAAGATCTGCGCTGAGAGCTCATTTTGGTCCCGAGGCGGCTGAGAT 1993
QY 1921 GTTATGCTGATTTGCTGAGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
| | | | |
Db 1994 GTTATGCTGATTTGCTGAGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2053
QY 1981 TTAAGAGCCGAGAGCTACTCCGATGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2040
| | | | |
Db 2054 TTAAGAGCCGAGAGCTACTCCGATGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2113
QY 2041 GACACTGTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGCGCTGTGAGG 2100
| | | | |
Db 2114 GACACTGTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGCGCTGTGAGG 2173
QY 2101 GCGTGGGCGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2160
| | | | |
Db 2174 GCGTGGGCGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2233
QY 2161 GAGGCTGACCTGCGAAGAGAGG 2184
| | | | |
Db 2234 GAGGCTGACCTGCGAAGAGAGG 2257

RESULT 5

PCT-US01-08631-8042
Sequence 8042, Application PC/TUS0108631
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 8042
LENGTH: 2652
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (136)..(1515)
OTHER INFORMATION: 29% homologous to Caenorhabditis elegans contains similarity
OTHER INFORMATION: to protein kinases (Pfam:PKinase_hmm, score: 149.36), accession
PCT-US01-08631-8042

Query Match 97.6% Score 2132.4; DB 1; Length 2652;
Best Local Similarity 99.4%; Pred. No. 0;

Matches 2172; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 1 ATGCTGAGGACCTGCGGGCCACCGATGTCATGAAAGACTTCTCGAGTTCGGGACGAG 60
DB 445 ATGCTGAGGACCTGCGGGCCACCGATGTCATGAAAGACTTCTCGAGTTCGGGACGAG 504
QY 61 GCCAGATCTCTGACGCGCTGACGACACCCCTGATGTCGCGCTCATCGGCATCAGCATC 120
DB 505 GCCAGATCTCTGACGCGCTGACGACACCCCTGATGTCGCGCTCATCGGCATCAGCATC 564
QY 121 CACCGCTCTGCTGCGCTGAGCTGCGCGCGCTCAGACGCTCAACACCGTCTGTC 180
DB 565 CACCGCTCTGCTGCGCTGAGCTGCGCGCGCTCAGACGCTCAACACCGTCTGTC 624
QY 181 GAGAACGCCAGAGATCTTCTTATACCTTGAGACATGCTCAACCAAAATAGCC 240
DB 625 GAGAACGCCAGAGATCTTCTTATACCTTGAGACATGCTCAACCAAAATAGCC 684
QY 241 TACCAATGCTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 685 TACCAATGCTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
QY 301 AAGTCGACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 745 AAGTCGACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 804
QY 361 TCTGACTACGGATTTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 805 TCTGACTACGGATTTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864
QY 421 CCGGGTACAGGCGCCAGAGATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 865 CCGGGTACAGGCGCCAGAGATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 924
QY 481 TTCTCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 925 TTCTCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 984
QY 541 CACCACTCCAGATTTGCAAGAAAGTGTCCAGAGGATCCCGGCTTCTGGGGCAGCG 600
DB 985 CACCACTCCAGATTTGCAAGAAAGTGTCCAGAGGATCCCGGCTTCTGGGGCAGCG 1044
QY 601 GAGGAGTGAAGTTCGGGCGACTGCGAGGCGCTGATGAGAGTGGGAGCTAAGCCA 660
DB 1045 GAGGAGTGAAGTTCGGGCGACTGCGAGGCGCTGATGAGAGTGGGAGCTAAGCCA 1104

QY 661 GAGAGCGACCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 1105 GAGAGCGACCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1164
QY 721 TTCATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 1165 TTCATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1224
QY 781 TACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 1225 TACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1284
QY 841 GAGAGGCGCTTATGAGTGCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 1285 GAGAGGCGCTTATGAGTGCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1344
QY 901 CTCAGGCTCAGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 1345 CTCAGGCTCAGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1404
QY 961 CTCAGGCTCAGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 1405 CTCAGGCTCAGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1464
QY 1021 ACCCTGCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1465 ACCCTGCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1524
QY 1081 GCCGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1525 GCCGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1584
QY 1141 TACCTGCTCAGACAGCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1585 TACCTGCTCAGACAGCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
QY 1201 CAGAACCCCTACCCAGATGAGGCGATGAGTGTGTCACAGCGGCTCTGAGTCTGCTGAC 1260
DB 1645 CAGAACCCCTACCCAGATGAGGCGATGAGTGTGTCACAGCGGCTCTGAGTCTGCTGAC 1704
QY 1261 AGCAATGGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1705 AGCAATGGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764
QY 1321 GAGCCCTACATGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 1765 GAGCCCTACATGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824
QY 1381 GAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1825 GAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1884
QY 1441 ACTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1885 ACTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
QY 1501 CCGGTGCGGCTTGGACAGGAAACCCGCGAGCCAGCCACACGCGCAACCCAAAGGTG 1560
DB 1945 CCGGTGCGGCTTGGACAGGAAACCCGCGAGCCAGCCACACGCGCAACCCAAAGGTG 2004
QY 1561 CCGGTGCGGCTTGGACAGGAAACCCGCGAGCCAGCCACACGCGCAACCCAAAGGTG 1620
DB 2005 CCGGTGCGGCTTGGACAGGAAACCCGCGAGCCAGCCACACGCGCAACCCAAAGGTG 2064
QY 1621 ATCTGATCAGCAGAAATCAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 2065 ATCTGATCAGCAGAAATCAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2124
QY 1681 CCACCCGCGAGGCTGCGAGGCTCCCTCAAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCT 1740
DB 2125 CCACCCGCGAGGCTGCGAGGCTCCCTCAAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCT 2184

QY 1741 AGTGTGCTTTCTCCACGACTGCGAGAGCTCAGACATGCTACATACGCGGCTCTCC 1800
 DB 2185 AGTGTGCTTTCTCCACGACTGCGAGAGCTCAGACATGCTACATACGCGGCTCTCC 2244
 QY 1801 TCCGACAGGTGTGACATGACCTGACCCCATGAGCGGGAGACCTTGACGACACCT - 1859
 DB 2245 TCCGACAGGTGTGACATGACCTGACCCCATGAGCGGGAGACCTTGACGACACCTG 2304
 QY 1860 GCAGGCGCTGAAG-ATCCTCGCGCTCAGAGACCTCATTTGGGTCCCGAGCGGCTGAG 1918
 DB 2305 GCAGGCGCTGAAGATCCTCGCGCTCAGAGACCTCATTTGGGTCCCGAGCGGCTGAG 2364
 QY 1919 ATGTATGTATGCTGCTGCTGAGAGATTTCTGAAGCCGAGCGGGCCGAGTCTATGCGC 1978
 DB 2365 ATGTATGTATGCTGCTGAGAGATTTCTGAGCGCCGAGCGGGCCGAGTCTATGCGC 2424
 QY 1979 TCTTAAAGCCGAGAGCTGACCTCGCATGGGCTGCTGATGCTCCCTGAGGCAA 2038
 DB 2425 CCTTAAAGCCGAGAGCTGACCTCGCATGGGCTGCTGATGCTCCCGAGGCAA 2484
 QY 2039 AGGACACTGTGTGTGC-ACCTTGAATGAATGAATGAGATGCTGCTGCTGCTG 2097
 DB 2485 AGGACACTGTGTGTGCAACCTTTGAAAAAGAAAAACAGAGTGTGCTGCTGCTGCTG 2544
 QY 2098 AGGGGCTGGGGGCGGAGGAGTGTGACATTTCTACAGCTCTACAGAGAGCTGGGCGG 2157
 DB 2545 AGGGGCTGGGGGCGGAGGAGTGTGACATTTCTACAGCTCTACAGAGAGCTGGGCGG 2604
 QY 2158 CTGAGGCTTGACACTGCGAAGAGAG 2183
 DB 2605 CTGAGGCTTGACACTGCGAAGAGAG 2630

RESULT 6
 US-10-132-382-5
 ; Sequence 5, Application US/10132382
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 7015
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-132-382-5

Query Match 88.7%; Score 1936.8; DB 40; Length 7015;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1;

QY 1 ATGCTGAGGACACTGCGGGGCCACCGCATGCGAAGAACTTCTCGAATTCGCGGACGAG 60
 DB 4124 ATGCTGAGGACACTGCGGGGCCACCGCATGCGAAGAACTTCTCGAATTCGCGGACGAG 4183
 QY 61 GCCAGCATGCTGACGCGCTGCGAGACCCCTGACATGCTGGGGCTCATGGGATCGACATC 120
 DB 4184 GCCAGCATGCTGACGCGCTGCGAGACCCCTGACATGCTGGGGCTCATGGGATCGACATC 4243
 QY 121 CACCGCTCTGCTGCGCTGAGAGCTGCGCGCTGACAGACCTCAACACGCTGCTGCTC 180
 DB 4244 CACCGCTCTGCTGCGCTGAGAGCTGCGCGCTGACAGACCTCAACACGCTGCTGCTC 4303
 QY 181 GAGAAAGCGCAAGATTTCTTCTTATACCTCTGGGACACATGCTACCCAAAAATAGCC 240
 DB 4304 GAGAAAGCGCAAGATTTCTTCTTATACCTCTGGGACACATGCTACCCAAAAATAGCC 4363
 QY 241 TACCATGATGCTGCGGCTGAGCTTACCTGACAGAAAGAAACATCATCTTCTGTAAGCTG 300
 DB 4364 TACCATGATGCTGCGGCTGAGCTTACCTGACAGAAAGAAACATCATCTTCTGTAAGCTG 4423

QY 301 AAGTGGACAAACATTTGTGTGTGTGCTCCCTTGACGTCAAGGACATCAACATCAAGCTA 360
 DB 4424 AAGTGGACAAACATTTGTGTGTGTGCTCCCTTGACGTCAAGGACATCAACATCAAGCTA 4483
 QY 361 TCTGACTACGGGATTTTGAGGACATTCATTCATGAGGCGCCCTAGGCGTGGAGGACT 420
 DB 4484 TCTGACTACGGGATTTTGAGGACATTCATTCATGAGGCGCCCTAGGCGTGGAGGACT 4543
 QY 421 CCTGGCTACAGGCGCCAGAGATAGAGCTTCGATTTGATATGATGAGAGATGATATG 480
 DB 4544 CCTGGCTACAGGCGCCAGAGATAGAGCTTCGATTTGATATGATGAGAGATGATATG 4603
 QY 481 TTTCTATGGAATGTGCTCTTACAGTGTCTGTACAGACAGCGCCCTGACACTGGGCGAC 540
 DB 4604 TTTCTATGGAATGTGCTCTTACAGTGTCTGTACAGACAGCGCCCTGACACTGGGCGAC 4663
 QY 541 CACGAGCTCCAGATTGGCAAGAACTGTCGAAGGCGCATCCGCGCGGTTCTGGGCGACCG 600
 DB 4664 CACGAGCTCCAGATTGGCAAGAACTGTCGAAGGCGCATCCGCGCGGTTCTGGGCGACCG 4723
 QY 601 GAGGAAGTGCAGTTCGCGGCGACTGACAGCGCTCATGATGAGTCTGAGGACATAAGCCA 660
 DB 4724 GAGGAAGTGCAGTTCGCGGCGACTGACAGCGCTCATGATGAGTCTGAGGACATAAGCCA 4783
 QY 661 GAGGAAGCGACCGCTGCGCTGTGCTGCTGAGTGAAGGAGCCGACTTTTGCCACC 720
 DB 4784 GAGGAAGCGACCGCTGCGCTGTGCTGCTGAGTGAAGGAGCCGACTTTTGCCACC 4843
 QY 721 TTTCTATGATGAATGCTGCTGTGAGGAAAGCAGACCTTCTCTCATCCAGGCGCAGAG 780
 DB 4844 TTTCTATGATGAATGCTGCTGTGAGGAAAGCAGACCTTCTCTCATCCAGGCGCAGAG 4903
 QY 781 TACACCGTGTGTGTTTGGAGATGAGAAAGAGAGTGCAGAACTACAGGCTGTGAACACA 840
 DB 4904 TACACCGTGTGTGTTTGGAGATGAGAAAGAGAGTGCAGAACTACAGGCTGTGAACACA 4963
 QY 841 GAGAAAGGCGCTCATGAGAGTGCAGAGATGTGCTGCCCTGGAGTGAAGTGAAGTGCACG 900
 DB 4964 GAGAAAGGCGCTCATGAGAGTGCAGAGATGTGCTGCCCTGGAGTGAAGTGAAGTGCACG 5023
 QY 901 CTCCAGGCTCCAGATCCCTGTGAGACGCCGAGGACCGAAGAAATCTACATCTAACCC 960
 DB 5024 CTCCAGGCTCCAGATCCCTGTGAGACGCCGAGGACCGAAGAAATCTACATCTAACCC 5057
 QY 961 CTCAGGCGATGTGCCCTTAAACACACCCACAGCGCTTGATCTCAGCTGTGCTC 1020
 DB 5058 CTCAGGCGATGTGCCCTTAAACACACCCACAGCGCTTGATCTCAGCTGTGCTC 5057
 QY 1021 ACCTGCTTCTTGGCGGCTGCTGTTATTAAGAAATCTACCTGATCTTACGCGGCGCTC 1080
 DB 5058 ACCTGCTTCTTGGCGGCTGCTGTTATTAAGAAATCTACCTGATCTTACGCGGCGCTC 5086
 QY 1081 GCCATGAGGCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1140
 DB 5087 GCCATGAGGCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 5146
 QY 1141 TACCTGTGCTCAGACACAGCCACAGGCTCAAGTTCAGCATCGGGGATGGAAGAGCGACGG 1200
 DB 5147 TACCTGTGCTCAGACACAGCCACAGGCTCAAGTTCAGCATCGGGGATGGAAGAGCGACGG 5206
 QY 1201 CAGAAACCCCTACCCAGTGAAGGCCATGAGTGTGCAACAGCGGCTCTGAGTGTGTATC 1260
 DB 5207 CAGAAACCCCTACCCAGTGAAGGCCATGAGTGTGCAACAGCGGCTCTGAGTGTGTATC 5266
 QY 1261 AGCAATGAGGCGGCGCTCTTGTATGATGACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 DB 5267 AGCAATGAGGCGGCGCTCTTGTATGATGACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 5326
 QY 1321 GAGGCTTACATGAGCGCCCTCCATGATGCTACGTCAGTCTGCTGCTGCTGCTGCTGCTG 1380
 DB 5327 GAGGCTTACATGAGCGCCCTCCATGATGCTACGTCAGTCTGCTGCTGCTGCTGCTGCTG 5386

QY 1381 GAGAGGTGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1440
 DB 5387 GAGAGGTGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5446
 QY 1441 ACCATACAGCTGTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1500
 DB 5447 ACCATACAGCTGTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5506
 QY 1501 CCCGTGCGGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1560
 DB 5507 CCCGTGCGGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5566
 QY 1561 CCTGAGGGGAGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1620
 DB 5567 CCTGAGGGGAGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5626
 QY 1621 ATCTGATGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1680
 DB 5627 ATCTGATGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5686
 QY 1681 CCAGCCGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1740
 DB 5687 CCAGCCGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5746
 QY 1741 AGTGTGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1800
 DB 5747 AGTGTGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5806
 QY 1801 TCCGACAGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1860
 DB 5807 TCCGACAGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5866
 QY 1861 CAGCGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1920
 DB 5867 CAGCGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5926
 QY 1921 GTATGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 1980
 DB 5927 GTATGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 5986
 QY 1981 TTTAAAGCCGAGAGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 2040
 DB 5987 TTTAAAGCCGAGAGCTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 6046
 QY 2041 GATAGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 2100
 DB 6047 GATAGCTGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 6106
 QY 2101 GCTGAGGGGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 2160
 DB 6107 GCTGAGGGGCTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 6166
 QY 2161 GAGCGTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 2184
 DB 6167 GAGCGTGTGCTGCTGATGACAGGCCACTCTGTGTGTATGACATCTCAC 6190

RESULT 7
 US-10-132-382-1
 ; Sequence 1, Application US/10132382
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 7093
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-132-382-1

Query Match. 88.7%; Score 1936.8; DB: 40; Length 7093;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1;

QY 1 ATCTGAGGACCTGCGGGCCACCGATGCTGAGAAAGACTTCTCCGAGTTCGGCAGAG 60
 DB 4202 ATCTGAGGACCTGCGGGCCACCGATGCTGAGAAAGACTTCTCCGAGTTCGGCAGAG 4261
 QY 61 GCCAGATGCTGAGCGGTGAGAGACCCCTGATGCTGAGGCTATGCGATCAGCATC 120
 DB 4262 GCCAGATGCTGAGCGGTGAGAGACCCCTGATGCTGAGGCTATGCGATCAGCATC 4321
 QY 121 CACCGCTGCTGCTGCTGAGAGCTGCGCGCTGAGAGCTGCTGAGAGCTGCTGCTG 180
 DB 4322 CACCGCTGCTGCTGCTGAGAGCTGCGCGCTGAGAGCTGCTGAGAGCTGCTGCTG 4381
 QY 181 GAGAGCGCAGAGATCTTCTTATACCTCTGAGAGACATGCTACCCAAAAATAGCC 240
 DB 4382 GAGAGCGCAGAGATCTTCTTATACCTCTGAGAGACATGCTACCCAAAAATAGCC 4441
 QY 241 TACAGATGCTGCGGGCTGAGAGCTGCTGAGAGAAATATCTTCTGAGCTG 300
 DB 4442 TACAGATGCTGCGGGCTGAGAGCTGCTGAGAGAAATATCTTCTGAGCTG 4501
 QY 301 AAGTCGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 DB 4502 AAGTCGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4561
 QY 361 TCTGACTAGGGATTTGAGAGCAGTATTCATGAGAGCGCCCTAGCGCTGAGGCACT 420
 DB 4562 TCTGACTAGGGATTTGAGAGCAGTATTCATGAGAGCGCCCTAGCGCTGAGGCACT 4621
 QY 421 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 4622 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4681
 QY 481 TTCTCTATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 4682 TTCTCTATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4741
 QY 541 CACACGCTCAGATTTGCCAAGAAAGTGTCCAGAGGCTCCGCCGTTCTGGGGCAGCG 600
 DB 4742 CACACGCTCAGATTTGCCAAGAAAGTGTCCAGAGGCTCCGCCGTTCTGGGGCAGCG 4801
 QY 601 GAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 DB 4802 GAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4861
 QY 661 GAGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 4862 GAGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4921
 QY 721 TTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 DB 4922 TTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4981
 QY 781 TACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 4982 TACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5041
 QY 841 GAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 DB 5042 GAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5101
 QY 901 CTCAGGCTCAGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 5102 CTCAGGCTCAGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5135
 QY 961 CTCAGGCTCAGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB 5136 CTCAGGCTCAGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5135

QY 1021 ACCTGCTTCTGCGCGCTGCTTATTAAGAAATTCCTACCTGCTGTAGCGGCGCTC 1080
DB 5136 -----AGAAATTCCTACCTGCTGTAGCGGCGCTC 5164
QY 1081 GCCGATGGGCTTGTGGCTGTGTTCCCGTGGGCGGACCCCAAGAGACAGCTGCTCC 1140
DB 5165 GCCGATGGGCTTGTGGCTGTGTTCCCGTGGGCGGACCCCAAGAGACAGCTGCTCC 5224
QY 1141 TACCTGTGCTCACAACACAGCCACAGGTCCAAAGTTCAGCATCGCGGATGAAGACGACG 1200
DB 5225 TACCTGTGCTCACAACACAGCCACAGGTCCAAAGTTCAGCATCGCGGATGAAGACGACG 5284
QY 1201 CAGAACCCCTACCCAGTAAGAGCCATGAGGTGTCACACAGCGCTGAGAGTCTGCTAC 1260
DB 5285 CAGAACCCCTACCCAGTAAGAGCCATGAGGTGTCACACAGCGCTGAGAGTCTGCTAC 5344
QY 1261 AGCAATGGGCGGCGCTGCTGTATGACATGCTGCTCCGAGAGATGTGAGGCGGCTG 1320
DB 5345 AGCAATGGGCGGCGCTGCTGTATGACATGCTGCTCCGAGAGATGTGAGGCGGCTG 5404
QY 1321 GAGCCCTACATGCGCCCTCCATGATGATGCTAGTGTGCTGAGTGTGAGGCGAGAGG 1380
DB 5405 GAGCCCTACATGCGCCCTCCATGATGATGCTAGTGTGCTGAGTGTGAGGCGAGAGG 5464
QY 1381 GAGAGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 5465 GAGAGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5524
QY 1441 ACCTACAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 5525 ACCTACAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5584
QY 1501 CCCGTGCGGCGCTTGTGACAGGAAACCCCGGACAGCCACAGCGCCCAACCAAGGTG 1560
DB 5585 CCCGTGCGGCGCTTGTGACAGGAAACCCCGGACAGCCACAGCGCCCAACCAAGGTG 5644
QY 1561 CCTGAGGGGAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB 5645 CCTGAGGGGAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5704
QY 1621 ATCTGATTCACACAGAGATTCATCTACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 5705 ATCTGATTCACACAGAGATTCATCTACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 5764
QY 1681 CCACCCCGCAGGCTGCCAGGTCCCTCAAGCTCCCAAGCTCCCAAGTCTTCC 1740
DB 5765 CCACCCCGCAGGCTGCCAGGTCCCTCAAGCTCCCAAGCTCCCAAGTCTTCC 5824
QY 1741 AGTGTGCTTCTCTCCAGGCTGCGAGACTGACATGCTACATAGCCCGGTGCTGCC 1800
DB 5825 AGTGTGCTTCTCTCCAGGCTGCGAGACTGACATGCTACATAGCCCGGTGCTGCC 5884
QY 1801 TCCGACAGGTGAGCATGACATGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 1860
DB 5885 TCCGACAGGTGAGCATGACATGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 5944
QY 1861 CAGGCGGTGAAGATCTCGCGCTCAGAGACTCATTTGGTCCCAAGCGCGGTGAGAT 1920
DB 5945 CAGGCGGTGAAGATCTCGCGCTCAGAGACTCATTTGGTCCCAAGCGCGGTGAGAT 6004
QY 1921 GTTATCGTATTTGGCTTGGAGAAAGATTTCTGAAGCCCAAGCGGCGGAGTCTGCTGCTG 1980
DB 6005 GTTATCGTATTTGGCTTGGAGAAAGATTTCTGAAGCCCAAGCGGCGGAGTCTGCTGCTG 6064
QY 1981 TTAAGAGCCGAGAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
DB 6065 TTAAGAGCCGAGAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6124
QY 2041 GACACTGTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGCTGCTGCTGCTGCTG 2100
DB 6125 GACACTGTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGCTGCTGCTGCTGCTG 6184
QY 2101 GGCTGGGCGCGCAGGAGTGTGACATTTTCAACAGTCTCTACGAGAGACTGGGCGGCTG 2160

DB 6185 GGCTGGGCGCGCAGGAGTGTGACATTTTCAACAGTCTCTACGAGAGACTGGGCGGCTG 6244
QY 2161 GAGGCTTGCACTCGCAAGAGAG 2184
DB 6245 GAGGCTTGCACTCGCAAGAGAG 6268

RESULT 8
US-10-132-382-7
Sequence 7, Application US/10132382
GENERAL INFORMATION:
APPLICANT: WEISS, BERTRAM
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
FILE REFERENCE: SCH-1811
CURRENT APPLICATION NUMBER: US/10/132.382
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 7229
TYPE: DNA
ORGANISM: Homo sapiens
US-10-132-382-7

Query Match 88.7%; Score 1936.8; DB 40; Length 7229;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1;

QY 1 ATGCTGAGGCACTTGGGCGCGCAGGATGCTGTAAGAACTTCTCGAGTTCCGCGAGAG 60
DB 4338 ATGCTGAGGCACTTGGGCGCGCAGGATGCTGTAAGAACTTCTCGAGTTCCGCGAGAG 4397
QY 61 GCCAGATGCTGACAGCGGCTGACAGCCCTGATCGTGGGCGCTCATCGGATCAGATC 120
DB 4398 GCCAGATGCTGACAGCGGCTGACAGCCCTGATCGTGGGCGCTCATCGGATCAGATC 4457
QY 121 CACCGGCTGCTGCTGCGCTGAGCTGCGCGGCTGACAGCCTCAACACCGTCTGCC 180
DB 4458 CACCGGCTGCTGCTGCGCTGAGCTGCGCGGCTGACAGCCTCAACACCGTCTGCC 4517
QY 181 GAGAACCCGAGATTTCTCTTATACCTCGGAGACATGCTACCCAAAAATATGCC 240
DB 4518 GAGAACCCGAGATTTCTCTTATACCTCGGAGACATGCTACCCAAAAATATGCC 4577
QY 241 TACAGATGCGGCTGCGGCTGAGCTGACCTGACAGAAACATCATCTTGTGACCTG 300
DB 4578 TACAGATGCGGCTGCGGCTGAGCTGACCTGACAGAAACATCATCTTGTGACCTG 4637
QY 301 AAGTGGCAACATTTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 4638 AAGTGGCAACATTTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4697
QY 361 TCTGACTACGGGATTTTGAGGCACTCATTCATGAGGCGGCTTATGAGTGAAGATGATG 420
DB 4698 TCTGACTACGGGATTTTGAGGCACTCATTCATGAGGCGGCTTATGAGTGAAGATGATG 4757
QY 421 CCTGGTACACAGGCGCCAGAGATCAGGCTGCTGATGATGATGATGATGATGATGATG 480
DB 4758 CCTGGTACACAGGCGCCAGAGATCAGGCTGCTGATGATGATGATGATGATGATGATG 4817
QY 481 TTTCTATGAGATGTGCTGTACAGTGTGCTGTACAGAGCGCCCTGCACTGGGCGAC 540
DB 4818 TTTCTATGAGATGTGCTGTACAGTGTGCTGTACAGAGCGCCCTGCACTGGGCGAC 4877
QY 541 CACAGGCTCCAGATTTGCAAGAGCTGCTCAAGGCGATCCCGGCTTGGGCGACCG 600
DB 4878 CACAGGCTCCAGATTTGCAAGAGCTGCTCAAGGCGATCCCGGCTTGGGCGACCG 4937
QY 601 GAGGAAGTGAAGTTCGCGGCTGCTGAGGCGCTCATGATGATGATGATGATGATGATG 660
DB 4938 GAGGAAGTGAAGTTCGCGGCTGCTGAGGCGCTCATGATGATGATGATGATGATGATG 4997

```

OY 661 GAGAGCAGCGCTGGCCCTGTGGTGTGAGCCAGATGAAGACCCGACTTTTGGCACC 720
    |||
DB 4998 GAGAGCAGCGCTGGCCCTGTGGTGTGAGCCAGATGAAGACCCGACTTTTGGCACC 5057
OY 721 TTGATGTATGAATCTGTGTGGGAGACAGACAGCCCTTCTCATGCCAGGCCAGGAG 780
    |||
DB 5058 TTGATGTATGAATCTGTGTGGGAGACAGACAGCCCTTCTCATGCCAGGCCAGGAG 5117
OY 781 TTGACCGGTGTGTGGGATGGAAGAGAGTCCAGGAATTAACAGGGTGTGAACACA 840
    |||
DB 5118 TTGACCGGTGTGTGGGATGGAAGAGAGTCCAGGAATTAACAGGGTGTGAACACA 5177
OY 841 GAGAGAGGCGCTCATGAGAGTGCAGAGAGATGCTGCTGGGATGAAGGTGAGCTGCCAG 900
    |||
DB 5178 GAGAGAGGCGCTCATGAGAGTGCAGAGAGATGCTGCTGGGATGAAGGTGAGCTGCCAG 5237
OY 901 CTCGAGGTCCAGAGATCCCTGTGGAGAGCTCCAGAGACCAGAAAATCTATCTACACC 960
    |||
DB 5238 CTCGAGGTCCAGAGATCCCTGTGGAGAGCTCCAGAGACCAGAAAATCTATCTACACC 5271
OY 961 CTCAGGGGATGTGCCCTTAACACACACCCCAAGGCCCTTGATCTCAGCTGTCTC 1020
    |||
DB 5272 ----- 5271
OY 1021 ACCTGCTTCTGGCCGTGCTGTATTAAAGAAATTCCTACCTGTCTAGCGGCTC 1080
    |||
DB 5272 -----AGAAATTCCTACCTGTCTAGCGGCTC 5300
OY 1081 GCCGATGGCTGTGTGGTGTGTTCCCTGTGGTGGGAGACCCCAAGAGACAGCTCTCC 1140
    |||
DB 5301 GCCGATGGCTGTGTGGTGTGTTCCCTGTGGTGGGAGACCCCAAGAGAGACAGCTCTCC 5360
OY 1141 TACCTGTCTCAGACAGACCAAGAGTCCAGATGCGGGATGAAGAGCGACGG 1200
    |||
DB 5361 TACCTGTCTCAGACAGACCAAGAGTCCAGATGCGGGATGAAGAGCGACGG 5420
OY 1201 CAGAACCCCTTACCAGATGAAGGCGATGAGTGTGTCAAGAGCGCTCTGAGTCTGTAC 1260
    |||
DB 5421 CAGAACCCCTTACCAGATGAAGGCGATGAGTGTGTCAAGAGCGCTCTGAGTCTGTAC 5480
OY 1261 AGCAATGGGCGGGGCTCTGTGTGATGAGTGTGCTCTCTGAGATCTGAGGCGGCTG 1320
    |||
DB 5481 AGCAATGGGCGGGGCTCTGTGTGATGAGTGTGCTCTCTGAGATCTGAGGCGGCTG 5540
OY 1321 GAGCCCTTACATGGCCCTTCCATGATGTTAGTCAAGTGTGTGAGTGTGAGGCGAGAG 1380
    |||
DB 5541 GAGCCCTTACATGGCCCTTCCATGATGTTAGTCAAGTGTGTGAGTGTGAGGCGAGAG 5600
OY 1381 GAGGAGGTGTGTGTGCTGTGATGACAAAGCCAACTCTTGTGTGATGATCACTCCACC 1440
    |||
DB 5601 GAGGAGGTGTGTGTGCTGTGATGACAAAGCCAACTCTTGTGTGATGATCACTCCACC 5660
OY 1441 ACCTTACAGTGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1500
    |||
DB 5661 ACCTTACAGTGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 5720
OY 1501 CCGGTGCGGCGCTTGGGAGACAGGAAACCCCGGAGCCAGCCAGAGGCGCAACCAAGGTG 1560
    |||
DB 5721 CCGGTGCGGCGCTTGGGAGACAGGAAACCCCGGAGCCAGCCAGAGGCGCAACCAAGGTG 5780
OY 1561 CCGGTGCGGCGCTTGGGAGACAGGAGTGTGAGATGATGATGATGATGATGATGATGATG 1620
    |||
DB 5781 CCGGTGCGGCGCTTGGGAGACAGGAGTGTGAGATGATGATGATGATGATGATGATGATG 5840
OY 1621 ATCTGTATCAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
    |||
DB 5841 ATCTGTATCAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 5900
OY 1681 CCAGCCCGCAGGCTGAGAGTCCCTCAAGGCTCCAGGCTCCAGAGAGGTTCTTCC 1740
    |||
DB 5901 CCAGCCCGCAGGCTGAGAGTCCCTCAAGGCTCCAGGCTCCAGAGAGGTTCTTCC 5960
OY 1741 AGTGTGCTTCTTCCAGCAGTGTGAGAGTGTGAGATGATGATGATGATGATGATGATG 1800

```

```

DB 5961 AGTGTGCTTCTTCCAGCAGTGTGAGAGTGTGAGATGATGATGATGATGATGATGATG 6020
OY 1801 TTCCAGAGGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
DB 6021 TTCCAGAGGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6080
OY 1861 CAGGCGGTGAGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
    |||
DB 6081 CAGGCGGTGAGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 6140
OY 1921 GTTATGCTATGAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1980
    |||
DB 6141 GTTATGCTATGAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 6200
OY 1981 TTTAAAGCCCGAGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATG 2040
    |||
DB 6201 TTTAAAGCCCGAGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATG 6260
OY 2041 GACACTGTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGAGCTGTGAGAG 2100
    |||
DB 6261 GACACTGTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGAGCTGTGAGAG 6320
OY 2101 GGCTGGGGGCGCAGGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATG 2160
    |||
DB 6321 GGCTGGGGGCGCAGGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATG 6380
OY 2161 GAGGCTGTGACCTGCAAGAGAGG 2184
    |||
DB 6381 GAGGCTGTGACCTGCAAGAGAGG 6404

```

RESULT 9
 US-10-132-382-3
 ; Sequence 3, Application US/10132382
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 7307
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-132-382-3

Query Match 88.7%; Score 1936.8; DB 40; Length 7307;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1;

```

OY 1 ATCTGAGGCACTGCGGGGCGCACCGATGCGATGAAGAACTTCTCCAGATTCGGCGAGAG 60
    |||
DB 4416 ATCTGAGGCACTGCGGGGCGCACCGATGCGATGAAGAACTTCTCCAGATTCGGCGAGAG 4475
OY 61 GCCAGATGCTGAGCGGCTGAGAGACCGCTGATGATGATGATGATGATGATGATGATGATG 120
    |||
DB 4476 GCCAGATGCTGAGCGGCTGAGAGACCGCTGATGATGATGATGATGATGATGATGATGATG 4535
OY 121 CACCCGCTGTGCTGCGCCCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 180
    |||
DB 4536 CACCCGCTGTGCTGCGCCCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 4595
OY 181 GAGAGCGCAGAGATGCTTCTTATACCCCTGTGAGACATGCTCAACCCAAAAAATAGGC 240
    |||
DB 4596 GAGAGCGCAGAGATGCTTCTTATACCCCTGTGAGACATGCTCAACCCAAAAAATAGGC 4655
OY 241 TACAGATGCTGTGGGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 300
    |||
DB 4656 TACAGATGCTGTGGGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 4715

```

```

OY 301 AAGTGGACAACTTCTGTGTGTGCTTGCATGACGTAAGAGACATCAACATCAAGCTA 360
Db 4716 AAGTGGACAACTTCTGTGTGTGCTTGCATGACGTAAGAGACATCAACATCAAGCTA 4775
OY 361 TCTGACTACGGGATTTGAGGAGATTCATTCATGAGGGGCGGCTTACGGGCTGAGAGGACT 420
Db 4776 TCTGACTACGGGATTTGAGGAGATTCATTCATGAGGGGCGGCTTACGGGCTGAGAGGACT 4835
OY 421 CCTGGCTACAGGCGCCAGAGATCAGGCGCTGCATTTATATATGATGAGAGATGATATG 480
Db 4836 CCTGGCTACAGGCGCCAGAGATCAGGCGCTGCATTTATATGATGAGAGATGATATG 4895
OY 481 TTCTCTATGGAATGCTCTCTACAGTTGCTGTGACAGAGAGCGGCTGCACTGGGCGAC 540
Db 4896 TTCTCTATGGAATGCTCTCTACAGTTGCTGTGACAGAGAGCGGCTGCACTGGGCGAC 4955
OY 541 CACGAGCTCCAGATGTCGCAAGAGCTGTCGCAAGGACATCCGCGGCTCTGGGAGAGCG 600
Db 4956 CACGAGCTCCAGATGTCGCAAGAGCTGTCGCAAGGACATCCGCGGCTCTGGGAGAGCG 5015
OY 601 GAGAAATGACAGTTCGCGGCTGCTGAGAGGCTCATGATGATGAGTGGAGACTTAAGCA 660
Db 5016 GAGAAATGACAGTTCGCGGCTGCTGAGAGGCTCATGATGATGAGTGGAGACTTAAGCA 5075
OY 661 GAGAAAGCAAGCGCTGCGGCTGCTGAGAGGCTCATGATGATGAGTGGAGACTTAAGCA 720
Db 5076 GAGAAAGCAAGCGCTGCGGCTGCTGAGAGGCTCATGATGATGAGTGGAGACTTAAGCA 5135
OY 721 TTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 5136 TTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5195
OY 781 TACACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 5196 TACACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5255
OY 841 GAGAAAGGCTCTATGAGAGTGTGAGAGAGTGTGCTGCTGGGATGAGAGTGTGAGAG 900
Db 5256 GAGAAAGGCTCTATGAGAGTGTGAGAGAGTGTGCTGCTGGGATGAGAGTGTGAGAG 5315
OY 901 CTCCAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 960
Db 5316 CTCCAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 5349
OY 961 CTCAAGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 5350 CTCAAGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5349
OY 1021 ACCTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1080
Db 5350 ACCTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 5378
OY 1081 GCCGATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 5379 GCCGATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5438
OY 1141 TACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 5439 TACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5498
OY 1201 CAGAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 5499 CAGAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5558
OY 1261 AGCAATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 5559 AGCAATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5618
OY 1321 GAGCCCTACATGAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
Db 5619 GAGCCCTACATGAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5678
OY 1381 GAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440

```

```

Db 5679 GAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5738
OY 1441 ACCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 5739 ACCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5798
OY 1501 CCGGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
Db 5799 CCGGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5858
OY 1561 CCTGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Db 5859 CCTGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5918
OY 1621 ATCTGTATCCACAGGAATCTACTGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 5919 ATCTGTATCCACAGGAATCTACTGATGATGATGATGATGATGATGATGATGATGATGAT 5978
OY 1681 CCACCCGCGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
Db 5979 CCACCCGCGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6038
OY 1741 AGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
Db 6039 AGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6098
OY 1801 TCGACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
Db 6099 TCGACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6158
OY 1861 CAGGCGGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
Db 6159 CAGGCGGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6218
OY 1921 GTTATGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
Db 6219 GTTATGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6278
OY 1981 TTAAAGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
Db 6279 TTAAAGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6338
OY 2041 GACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
Db 6339 GACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6398
OY 2101 GCTGTGGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
Db 6399 GCTGTGGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6458
OY 2161 GAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
Db 6459 GAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6482

```

```

RESULT 10
PCT-US02-05109-456
Sequence 456, Application PC/TUS0205109
GENERAL INFORMATION:
APPLICANT: Hyseq Inc.
APPLICANT: Tang, Tom Y.
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, RyLe
APPLICANT: Asundi, Vinod.
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Ma, Yundong
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong

```

[illegible]

```

; LIFE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (872)..(1831)
US-09-810-173-456

```

Query Match 49.88; Score 1088.2; DB 31; Length 2839;
 Best Local Similarity 97.18; Pred. No. 6,6e-214;
 Matches 1119; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 1032 GGGCGTGGCTTTTAAAGAAATTCCTACCTGCTCTTACGGGGCTCGCCGATGGCT 1091
 DB 680 GGTCTTCACACTGTGTGACAGAAAGGTGACCTGTCTTACG-GGGCTGGCGATGGCT 738
 QY 1092 TGTGGCTGTGTTCCGTTGGCGGGGACCCCAAGAGACAGTGTCTACCTGTGCTC 1151
 DB 739 TGTGTGTGTTCCCGTGTGCGGGGACCCCAAGAGACAGTGTCTCTCTGTGCTC 798
 QY 1152 ACACAGACAGCAAGTTCAGATCGCGATGAGAGCAGCAGCAAGCCCTTA 1211
 DB 799 ACACAGACAGCAAGTTCAGATCGCGATGAGAGCAGCAGCAAGCCCTTA 858
 QY 1212 CCCAGTGAAGGCGATGAGAGGCGTCAACAGCGCTGTGAGTGTGTACACATGGGCC 1271
 DB 859 CCCAGTGAAGGCGATGAGAGGCGTCAACAGCGCTGTGAGTGTGTACACATGGGCC 918
 QY 1272 GGGCTCTCTTCATGCACTGTGCTCCCTGAGATGTGAGGGGCTGAGCCCTACAT 1331
 DB 919 GGGCTCTCTTCATGCACTGTGCTCCCTGAGATGTGAGGGGCTGAGCCCTACAT 978
 QY 1332 GGGCCCTCATGCTTACGTCAGTGTGTGAGCTGTGAGGGCAGAGGAGGTGT 1391
 DB 979 GGGCCCTCATGCTTACGTCAGTGTGTGAGCTGTGAGGGCAGAGGAGGTGT 1038
 QY 1392 CTGTGCTGTGATGCAAGGCAACTCTGTGTATGATGATGATGATGATGATGATGAT 1451
 DB 1039 CTGTGCTGTGATGCAAGGCAACTCTGTGTATGATGATGATGATGATGATGATGAT 1098
 QY 1452 GTGTGCGGCTACTTGTGCGGGGTCCCGAGCCCTCAGGAGCAATGTTCCGCGGCC 1511
 DB 1099 GTGTGCGGCTACTTGTGCGGGGTCCCGAGCCCTCAGGAGCAATGTTCCGCGGCC 1158
 QY 1512 CTGTGACACAGGACCCCGGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1571
 DB 1159 CTGTGACACAGGACCCCGGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1218
 QY 1572 CTCATGCGGAGCGTGAAGCATGATGATGATGATGATGATGATGATGATGATGAT 1631
 DB 1219 CTCATGCGGAGCGTGAAGCATGATGATGATGATGATGATGATGATGATGATGAT 1278
 QY 1632 CCAGAAATCACTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1691
 DB 1279 CCAGAAATCACTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1338
 QY 1692 GGTGCGGAGTCCCTCAGCCCTCCAGCTCCCGAGCAAGTCTTCCAGTGTGCTTT 1751
 DB 1339 GGTGCGGAGTCCCTCAGCCCTCCAGCTCCCGAGCAAGTCTTCCAGTGTGCTTT 1398
 QY 1752 CTCACGAGCTGGAGGAGCTGAGACATGCTACATAGCGCGGCTGTGCTCCGACGTC 1811
 DB 1399 CTCACGAGCTGGAGGAGCTGAGACATGCTACATAGCGCGGCTGTGCTCCGACGTC 1458
 QY 1812 TGACATGACCTGACCCCATGAGAGGGGAGACCTTACGACAGCCTGAGGGCGTAA 1871
 DB 1459 TGACATGACCTGACCCCATGAGAGGGGAGACCTTACGACAGCCTGAGGGCGTAA 1518
 QY 1872 GATGCTGCGCTGAGAGACTCATTTGGGTCCCAAGGGCGGTGAGATGTTATGCTAT 1931
 DB 1519 GATGCTGCGCTGAGAGACTCATTTGGGTCCCAAGGGCGGTGAGATGTTATGCTAT 1578
 QY 1932 TGGGCTGAGAGAGATTTGTAAGCCAGGGGGCGGAGTCAATTTGCCCTTAAAGCCCG 1991
 DB 1579 TGGGCTGAGAGAGATTTGTAAGCCAGGGGGCGGAGTCAATTTGCCCTTAAAGCCCG 1638
 QY 1992 AGAGCTGACTCCGATGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2051
 DB 1639 AGAGCTGACTCCGATGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1698
 QY 2052 GTGCACTTTGAAAAATGAAAAACAGAGAGTGTGCTGCGCTGTGAGGGGCTGGGGCGC 2111

DB 1699 GTGCACTTTGAAAAATGAAAAACAGAGTGTGCTGCGCTGAGAGGGCTGGGGCGC 1758
 QY 2112 CAGGAGTTTGACATTTTTCACAGTCTCTACAGAGACTGGCGGCTGAGGCTTACAC 2171
 DB 1759 CAGGAGTTTGACATTTTTCACAGTCTCTACAGAGACTGGCGGCTGAGGCTTACAC 1818
 QY 2172 TCGCAGAGAGAG 2184
 DB 1819 TCGCAGAGAGAG 1831

RESULT 12
 US-09-471-275-7370
 ; Sequence 7370, Application US/09471275
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: Novel Contigs Obtained
 ; FILE REFERENCE: 782
 ; FILE REFERENCE: From various Libraries
 ; CURRENT APPLICATION NUMBER: US/09/471,275
 ; EARLIER FILING DATE: 1999-12-23
 ; EARLIER APPLICATION NUMBER: US 09/235,076
 ; EARLIER FILING DATE: 1999-01-20
 ; EARLIER APPLICATION NUMBER: US 09/234,611
 ; EARLIER FILING DATE: 1999-01-22
 ; EARLIER APPLICATION NUMBER: US 09/240,371
 ; EARLIER FILING DATE: 1999-01-29
 ; EARLIER APPLICATION NUMBER: US 09/277,227
 ; EARLIER FILING DATE: 1999-03-25
 ; EARLIER APPLICATION NUMBER: US 09/271,490
 ; EARLIER FILING DATE: 1999-03-18
 ; EARLIER APPLICATION NUMBER: US 09/293,972
 ; EARLIER FILING DATE: 1999-04-15
 ; EARLIER APPLICATION NUMBER: US 09/274,861
 ; EARLIER FILING DATE: 1999-03-23
 ; EARLIER APPLICATION NUMBER: US 60/125,453
 ; EARLIER FILING DATE: 1999-03-19
 ; EARLIER APPLICATION NUMBER: US 60/126,605
 ; EARLIER FILING DATE: 1999-03-26
 ; EARLIER APPLICATION NUMBER: US 09/306,350
 ; EARLIER FILING DATE: 1999-05-07
 ; EARLIER APPLICATION NUMBER: US 09/399,720
 ; EARLIER FILING DATE: 1999-09-21
 ; EARLIER APPLICATION NUMBER: US 09/404,284
 ; EARLIER FILING DATE: 1999-09-21
 ; EARLIER APPLICATION NUMBER:
 ; NUMBER OF SEQ ID NOS: 1999-12-16
 ; SOFTWARE: PL_CT_genes Version 1.0
 ; SEQ ID NO 7370
 ; LENGTH: 995
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (298) ... (811)
 ; OTHER INFORMATION: similar to g1530878 in the genepept database release 114,
 ; OTHER INFORMATION: Run with FASTX 3.3c00, default parameters
 US-09-471-275-7370

Query Match 43.18; Score 941.6; DB 18; Length 995;
 Best Local Similarity 98.08; Pred. No. 9,4e-184;
 Matches 964; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 1110 GTGCGGGGACCCCAAGAGACAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1169
 DB 12 GGTGCGGGGACCCCAATTTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 71
 QY 1170 CAAATTAGCATCCGGATGAGAGCGCAGGAGCAACCCCTACCCATGTAAGGCAATGGA 1229
 DB 72 CAAATTAGCATCCGGATGAGAGCGCAGGAGCAACCCCTACCCATGTAAGGCAATGGA 131


```
Db 13764 GTGCTCTCCCTGAGATCTGAGGGGCTGGAGCCCTACATGCCCCCTCCATGTTAGT 13823
QY 1352 CAGTCGCTGCTCAGCTCTGAGGGGAGAGGGGAGGTCGTGCTGCTGGATGACAGG 1411
Db 13824 CAGTCGCTGCTCAGCTCTGAGGGGAGAGGGGAGGTCGTGCTGCTGGATGACAGG 13883
QY 1412 CCAACTCTTGGTGTATGATACCACTCCACACTACAGTGTGTGCCGGTACTTCTGCG 1471
Db 13884 CCAACTCTTGGTGTATGATACCACTCCACACTACAGTGTGTGCCGGTACTTCTGCG 13943
QY 1472 GGGTCCCCAGCCCCCTCAGGGACATGTTCCCGCTGCGGCCCTTGACACGGAAACCCCGG 1531
Db 13944 GGGTCCCCAGCCCCCTCAGGGACATGTTCCCGCTGCGGCCCTTGACACGGAAACCCCGG 14003
QY 1532 CAGCCACCCACAGGGCCCAACCCAAAGGTGCTGAGGGGGAGCTCCATCGCGGAGCTGAGCA 1591
Db 14004 CAGCCACCCACAGGGCCCAACCCAAAGGTGCTGAGGGGGAGCTCCATCGCGGAGCTGAGCA 14063
QY 1592 TCATGTACAGTGTAGAGAGCTGGGACAGAGATCCTGATCCACAGAAATCACTACTGACT 1651
Db 14064 TCATGTACAGTGTAGAGAGCTGGGACAGAGATCCTGATCCACAGAAATCACTACTGACT 14123
QY 1652 ACTGCTCCATGTCCTCTCTACTCTCTATCCCAACCCGCGCAGGCTGCCAGGTCCCCCTCAA 1711
Db 14124 ACTGCTCCATGTCCTCTCTACTCTCTATCCCAACCCGCGCAGGCTGCCAGGTCCCCCTCAA 14183
QY 1712 GCCTCCCCAGCTCCCCAGCAAGTCTTCCAGTGTGCTTCTCCACCGACTGCGAGGACT 1771
Db 14184 GCCTCCCCAGCTCCCCAGCAAGTCTTCCAGTGTGCTTCTCCACCGACTGCGAGGACT 14243
QY 1772 CAGACATGCTACATAGCGCCGCTGCTGCTCCGACAGGTGTGAGCATGATGACCTGACCCCA 1831
Db 14244 CAGACATGCTACATAGCGCCGCTGCTGCTCCGACAGGTGTGAGCATGATGACCTGACCCCA 14303
QY 1832 TGGACGGGGAGACCTTGAGCCAGCAGCTGAGGCCGTGAAGATCCTCGCCCTCAGAGACC 1891
Db 14304 TGGACGGGGAGACCTTGAGCCAGCAGCTGAGGCCGTGAAGATCCTCGCCCTCAGAGACC 14363
QY 1892 TCATTTGGGTCCCAAGGCGCGGTGAGATGTTATCGTCATTTGGCTGAGAGAAGATTCTG 1951
Db 14364 TCATTTGGGTCCCAAGGCGCGGTGAGATGTTATCGTCATTTGGCTGAGAGAAGATTCTG 14423
QY 1952 AAGCCAGCGGGGGCGAG 1969
Db 14424 CTCTGCTCTGGGGACAG 14441
```

Search completed: April 15, 2003, 06:31:19
Job time : 3391 secs

OY	1230	GGGATGCAACAGGGGCTGAGAGGTCTGGACAGCAATGAGGCGGGGCTCTTGTCATCGA	1289
OY	1230	GGGATGCAACAGGGGCTGAGAGGTCTGGACAGCAATGAGGCGGGGCTCTTGTCATCGA	1289
Db	132	GGGATGCAACAGGGGCTGAGAGGTCTGGACAGCAATGAGGCGGGGCTCTTGTCATCGA	191
OY	1290	CTGAGCTCCCTCGAATATGACAGGGGCTGAGAGGCTTACATATGGGCGCCCTCATGTGTAC	1349
Db	192	CTGAGCTCCCTCGAATATGACAGGGGCTGAGAGGCTTACATATGGGCGCCCTCATGTGTAC	251
OY	1350	GTCAGTCGTGTGACGCTCTGAGAGGCGAGAGGGAGAGAGGTGCTTGTTGGCTGTGATGACAA	1409
Db	252	GTCAGTCGTGTGACGCTCTGAGAGGCGAGAGGGAGAGAGGTGCTTGTTGGCTGTGATGACAA	311
OY	1410	GGCCAACTCTTGTGTATGTACATCTCCACACACTTACACAGCTGTGTGCTGGGTACTTCTG	1469
Db	312	GGCCAACTCTTGTGTATGTACATCTCCACACACTTACACAGCTGTGTGCTGGGTACTTCTG	371
OY	1470	CGGGGTCCTCCAGGCGCCCTCGAGGACATGTTTCCGCGGCGCCCTTGGACACGGAACCCCC	1529
Db	372	CGGGGTCCTCCAGGCGCCCTCGAGGACATGTTTCCGCGGCGCCCTTGGACACGGAACCCCC	431
OY	1530	GGCAGCCAGCCACACAGGGCAACCCAAAGGTGCTGAGGGGAGCTCATCGCGGACGTGAG	1589
Db	432	GGCAGCCAGCCACACAGGGCAACCCAAAGGTGCTGAGGGGAGCTCATCGCGGAGGTGAG	491
OY	1590	CATATGTATGACGTGAGGAGCTGGGACACGACAGATCCTGATTCACACAGAAATCACTCATGTA	1649
Db	492	CATATGTATGACGTGAGGAGCTGGGACACGACAGATCCTGATTCACACAGAAATCACTCATGTA	551
OY	1650	CTATGCTCCATGTCTCTCCCTCACTCTCACTCACTCCACGCCCGGACAGGTGCCAGTCCCTCT	1709
Db	552	CTATGCTCCATGTCTCTCCCTCACTCTCACTCACTCCACGCCCGGACAGGTGCCAGTCCCTCT	611
OY	1710	AAGCCTCCCGGAGCCCGCCACACAGTCTTCAAGATGAGCTTCTCTCACAGACTGAGAGA	1769
Db	612	AAGCCTCCCGGAGCCCGCCACACAGTCTTCAAGATGAGCTTCTCTCACAGACTGAGAGA	671
OY	1770	CTCAGACATGCTATCATACGCCCGGTGCTGCTCCGACAGGTGTGAGCATGACCTTACCCC	1829
Db	672	CTCAGACATGCTATCATACGCCCGGTGCTGCTCCGACAGGTGTGAGCATGACCTTACCCC	731
OY	1830	CATGAGCGGGGAGACCTTCAAGCCAGCACTGCAAGCGCGGTGAAGATCTCTGCCGTGAGAGA	1889
Db	732	CATGAGCGGGGAGACCTTCAAGCCAGCACTGCAAGCGCGGTGAAGATCTCTGCCGTGAGAGA	791
OY	1890	CCTCATTTTGGGTCCCGAGGGCGGCGGTGGAGATGTTATCGATTTGGGCTGGAGAAAGATTTC	1949
Db	792	CCTCATTTTGGGTCCCGAGGGCGGCGGTGGAGATGTTATCGATTTGGGCTGGAGAAAGATTTC	851
OY	1950	TGAAGCCACGAGGGGCGGAGTCAATTTGCCCTTAAAGCCGACAGAGTACTCCCATGAG	2009
Db	852	TGAAGCCACGAGGGGCGGAGTCAATTTGCCCTTAAAGCCGACAGAGTACTCCCATGAG	911
OY	2010	GGTGCTGTGTGATCTCTCCGTGCTGTGCAAAAGACACTGTTGTGTGC -ACCTTTGAAAATG	2069
Db	912	GGTGCTGTGTGATCTCTCCGTGCTGTGCAAAAGACACTGTTGTGTGTGCAACTTTGAAAAG	971
OY	2069	AAAACACAGATGCTGCTGGCCG 2092	
Db	972	AAAACACAGATGCTGCTGGCCG 995	

```

? APPLICANT: Villevall, Jean-Luc M. G.
? APPLICANT: Cepeda, Mario
? APPLICANT: Kingsbury, Gillian A.
? APPLICANT: Busfield, Samantha J.
? TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
? TITLE OF INVENTION: THEREFOR
? FILE REFERENCE: 1600.1152-001
? CURRENT APPLICATION NUMBER: US/09/637,890
? CURRENT FILING DATE: 2000-08-09
? PRIOR APPLICATION NUMBER: 60/147,939
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 10217
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 9034
? LENGTH: 1912
? TYPE: DNA
? ORGANISM: Homo sapiens
? OS-09-637-890-9034

```


110 GCATCAGCATCCACCCGCTCTGCTTCGCCCTGGAGCTCGCGCCGCTCAGCAGCCTCAACA 165

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into to two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpn** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapn** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

Db 1 GCATCAGATCCACCCGCTGCTGCTCCGCTGAGACTCGCCGCTCAGCAGCTCAACA 60
QY 170 CCGGTCTGTCGAGAGAGCCAGAGATTCCTTATACCCCTGGAGACATGCTCAACC 229
Db 61 CCGGTCTGTCGAGAGAGCCAGAGATTCCTTATACCCCTGGAGACATGCTCAACC 120
QY 230 AAAAAATGCTTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 289
Db 121 AAAAAATGCTTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 290 TCTGAGCTGAGTGGAGACAATTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 349
Db 181 TCTGTGACCTTAAAGTGGAGACAATTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 350 ACATCAAGCTATCTGACTACGGGATTTGAGAGCATCTTCATGAGAGGCGCCCTAGGCG 409
Db 241 ACATCAAGCTATCTGACTACGGGATTTGAGAGCATCTTCATGAGAGGCGCCCTAGGCG 300
QY 410 TGGAGGCACTCTGCTGCTACAGGCGCCAGAGATCAGGCGCTGCGATTTATATGATGAGA 469
Db 301 TGGAGGCACTCTGCTGCTACAGGCGCCAGAGATCAGGCGCTGCGATTTATATGATGAGA 360
QY 470 AGGTAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 529
Db 361 AGGTAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 530 CACTGCGCCACACAGCTCCAGATTTGCAAGAGCTGTCCAGAGGCTCCGCGGCTTC 589
Db 421 CACTGCGCCACACAGCTCCAGATTTGCAAGAGCTGTCCAGAGGCTCCGCGGCTTC 480
QY 590 TGGGGGACCGGAGAGTGTAGTTCGGGCGACTGCGAGGCGCTATGATGAGTGTGCGG 649
Db 481 TGGGGGACCGGAGAGTGTAGTTCGGGCGACTGCGAGGCGCTATGATGAGTGTGCGG 540
QY 650 ACATTAAGCCAGAGAGAGGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709
Db 541 ACATTAAGCCAGAGAGAGGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 710 CTTTGGCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 769
Db 601 CTTTGGCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 770 AGGGCCAGGAGTACACCGT 829
Db 661 AGGGCCAGGAGTACACCGT 720
QY 830 TGTGTAACACAGAGAGAGGCTCATGAGTGTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 889
Db 721 TGTGTAACACAGAGAGAGGCTCATGAGTGTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 890 TGAAGCTGCACTCCAGTCCAGATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 949
Db 781 TGAAGCTGCACTCCAGTCCAGATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 950 ACATCTACACCTCAAGGGAGTGTGCTTAAACACACCCCAACAGGCTTGTGATATCTC 1009
Db 841 ACATCTACACCTCAAGGGAGTGTGCTTAAACACACCCCAACAGGCTTGTGATATCTC 900
QY 1010 CAGCTGTGCTACCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069
Db 901 CAGCTGTGCTACCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 1070 TAGGGGCGCTGCGGATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1129
Db 961 TAGGGGCGCTGCGGATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1130 ACAGCTGCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189
Db 1021 ACAGCTGCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1190 AAGAGCAGGCGAGAACCCCTACCCAGTGAAGGCGCATGAGTGTGCTGAACAGGCGCTCTG 1249
Db 1081 AAGAGCAGGCGAGAACCCCTACCCAGTGAAGGCGCATGAGTGTGCTGAACAGGCGCTCTG 1140

QY 1250 AGGTCTGTACAGCAATGGGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309
Db 1141 AGGTCTGTACAGCAATGGGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1310 GCAGGCGGCTGAGGCGCTTACATGAGCGCCCTTCATGATGTTACGTACGTGTGAGCTCTG 1369
Db 1201 GCAGGCGGCTGAGGCGCTTACATGAGCGCCCTTCATGATGTTACGTACGTGTGAGCTCTG 1260
QY 1370 AGGGCAGGAGGAGAGT 1429
Db 1261 AGGGCAGGAGGAGAGT 1320
QY 1430 ACCACTCCACCACTTACAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1489
Db 1321 ACCACTCCACCACTTACAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1490 GGGACATGTTTCCGTCGCGGCTTGGACACGGAAACCCCGGACGCCACACGGCCA 1549
Db 1381 GGGACATGTTTCCGTCGCGGCTTGGACACGGAAACCCCGGACGCCACACGGCCA 1440
QY 1550 ACCCAAGGTGCTGAGGGGAGCTCCATCGCGAGCTGAGCATCATCATCATCATCATCATCAT 1609
Db 1441 ACCCAAGGTGCTGAGGGGAGCTCCATCGCGAGCTGAGCATCATCATCATCATCATCATCAT 1500
QY 1610 TGGGACGAGATCTGATCCACAGAGAACTCACTGACTGACTGCTGCTGCTGCTGCTGCTGCT 1669
Db 1501 TGGGACGAGATCTGATCCACAGAGAACTCACTGACTGACTGCTGCTGCTGCTGCTGCTGCT 1560
QY 1670 ACTGCTATCCCAACCCCGCAGGCTGCGAGTCCCTTCAAGCTCCACGCTCCCAAG 1729
Db 1561 ACTGCTATCCCAACCCCGCAGGCTGCGAGTCCCTTCAAGCTCCCAAGCTCCCAAG 1620
QY 1730 CAAGTCTTCAGAGTGTGCTTTCACAGCACTGCGAGGAGCTGAGCATGCTCATATGCG 1789
Db 1621 CAAGTCTTCAGAGTGTGCTTTCACAGCACTGCGAGGAGCTGAGCATGCTCATATGCG 1680
QY 1790 CCGGTGCTGCTCCGACAGGCTGAGCATGACCTGACCCCATGAGAGGAGAGACCTTCA 1849
Db 1681 CCGGTGCTGCTCCGACAGGCTGAGCATGACCTGACCCCATGAGAGGAGAGACCTTCA 1740
QY 1850 GCCAGACCTGCGAGGCGGCTGAAGTCTGCGCGCTGACAGACCTCATTTGGCTCCCAAGG 1909
Db 1741 GCCAGACCTGCGAGGCGGCTGAAGTCTGCGCGCTGACAGACCTCATTTGGCTCCCAAGG 1800
QY 1910 GCGGTGAGATGTTATGCTGATGCTGAGAGAGATTTGAGAGCCACAGGCGGCGAG 1969
Db 1801 GCGGTGAGATGTTATGCTGATGCTGAGAGAGATTTGAGAGCCACAGGCGGCGAG 1860
QY 1970 TCATTTGCGCTTAAAGCCCGAGAGCTGACTCGCATG- 2008
Db 1861 TCATTTGCGCTTAAAGCCCGAGAGCTGACTCGCATG- 1920
QY 2009 - 2008
Db 1921 TGAAGTATGCTGGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
QY 2009 - 2008
Db 1981 GAGGCTGGCGCAGCTCAGCTCCCACTGTTTACACCCCGCAGCTTGTACACAGC 2040
QY 2009 - 2008
Db 2041 CCGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
QY 2062 GAAATTAAGAAACACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2121
Db 2101 GAAATTAAGAAACACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
QY 2122 GACATTTTACAGTCTCTGAGAGAGCTGGCGGCTGAGAGCTTGTGCTGCTGCTGCTGCTGCT 2181
Db 2161 GACATTTTACAGTCTCTGAGAGAGCTGGCGGCTGAGAGCTTGTGCTGCTGCTGCTGCTGCT 2220

OY 2182 AGG 2184
Db 2221 AGG 2223

RESULT 2

US-10-094-749-50
Sequence 50, Application US/10094749
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOTIRO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 50
LENGTH: 3112
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-50

Query Match 84.2% Score 1840: DB 8: Length 3112;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1851: Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 333 CGTCAAGGAGACATCAATCAAGTATCTGACTACGAGGATTTGAGGCGATCTTCCA 392
Db 1 CGTCAAGGAGACATCAATCAAGTATCTGACTACGAGGATTTGAGGCGATCTTCCA 60
OY 393 TGAAGGCGCCCTAGGCTGAGGAGGACTCTGCTTACAGGCCAGAGATCAGGCTCG 452
Db 61 TGAAGGCGCCCTAGGCTGAGGAGGACTCTGCTTACAGGCCAGAGATCAGGCTCG 120
OY 453 CATGTATATGAGAGAGAGATATGTCTCTATGGAATGCTGCTTACAGAGTGTCT 512
Db 121 CATGTATATGAGAGAGAGATATGTCTCTATGGAATGCTGCTTACAGAGTGTCT 180
OY 513 GTCAAGAGAGGCGCTGACTGAGGCGACACAGCTCAGATTGCGAAGAGCTGTCCAA 572
Db 181 GTCAAGAGAGGCGCTGACTGAGGCGACACAGCTCAGATTGCGAAGAGCTGTCCAA 240
OY 573 GGGCATCCCGCGGTTCTGGGGCAGCCGGAAGAGTGCAGTTCGGGCACTGACGGCT 632
Db 241 GGGCATCCCGCGGTTCTGGGGCAGCCGGAAGAGTGCAGTTCGGGCACTGACGGCT 300
OY 633 CATATGAGAGCTGGGAGCACTAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
Db 301 CATATGAGAGCTGGGAGCACTAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 693 CCAGATGAG 752
Db 1441 AGACATCTAATACAGCCGCGGTGCTGCTCCGACAGAGTGTGAGATGACCTGACCCCAT 1500

Db 361 CCAGATGAG 420
OY 753 AGCTTCTTCTCATCCAGAGGCGAGAGATACACCGATGTTTGGAGTGAAGAGAGAG 812
Db 421 AGCTTCTTCTCATCCAGAGGCGAGAGATACACCGATGTTTGGAGTGAAGAGAGAG 480
OY 813 GTCCAG 872
Db 481 GTCCAG 540
OY 873 CTGCCCTGGATGAG 932
Db 541 CTGCCCTGGATGAG 600
OY 933 CGAG 992
Db 601 CGAG 660
OY 993 ACAGGCTTGGATGAG 1052
Db 661 ACAGGCTTGGATGAG 720
OY 1053 GAATTCCTACCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1112
Db 721 GAATTCCTACCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
OY 1113 GCGGGGACCCCAAG 1172
Db 781 GCGGGGACCCCAAG 840
OY 1173 GTTCAGATGAG 1232
Db 841 GTTCAGATGAG 900
OY 1233 GTTCAGATGAG 1292
Db 901 GTTCAGATGAG 960
OY 1293 TGCCTCCTTGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1352
Db 961 TGCCTCCTTGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
OY 1353 AGTGTGTGAG 1412
Db 1021 AGTGTGTGAG 1080
OY 1413 CAATCTCTTGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
Db 1081 CAATCTCTTGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
OY 1473 GGTCCCAAGCCCTCAG 1532
Db 1141 GGTCCCAAGCCCTCAG 1200
OY 1533 AGCAGGACAG 1592
Db 1201 AGCAGGACAG 1260
OY 1593 CATGTATATGAG 1652
Db 1261 CATGTATATGAG 1320
OY 1653 CTGCTCCTATGATGATGAG 1712
Db 1321 CTGCTCCTATGATGATGAG 1380
OY 1713 CTTCCCGAGGCTCCAG 1772
Db 1381 CTTCCCGAGGCTCCAG 1440
OY 1773 AGACATCTAATACAGCCGCGGTGCTGCTCCGACAGAGTGTGAGATGACCTGACCCCAT 1832
Db 1441 AGACATCTAATACAGCCGCGGTGCTGCTCCGACAGAGTGTGAGATGACCTGACCCCAT 1500

QY	1833	GGAGGGGGAGACCTTTCAGCCAGACCTGAGAGCGCGGAAGATCTTCGCCGTAGAGACT	1892
Db	1501	GGAGGGGGAGACCTTTCAGCCAGACCTGAGAGCGCGGAAGATCTTCGCCGTAGAGACT	1560
QY	1893	CATTGGGTCGCCAGCGCGGTGGAGATGTTATCCGTCATTGGCCCTGAGAGAAAGATTCTGA	1952
Db	1561	CATTGGGTCGCCAGCGCGGTGGAGATGTTATCCGTCATTGGCCCTGAGAGAAAGATTCTCGA	1619
QY	1953	AGCCCAAGCGGGGCCGAGTCAATTGCCGTCTTAAAAAGCCCGAGAGCTGACTCCGATGGGT	2012
Db	1620	AGCCCAAGCGGGGCCGAGTCAATTGCCGTCTTAAAAAGCCCGAGAGCTGACTCCGATGGGT	1679
QY	2013	GCTGGTGGATGCTGCCTGGTGGCCAAAGACACTGTGTGTGCACCTTGTGAAATGAAAA	2072
Db	1680	GCTGGTGGATGCTGCCTGGTGGCCAAAGACACTGTGTGTGCACCTTGTGAAATGAAAA	1739
QY	2073	CACAGATGCTGCCTGCCTGTGAGAGGGCGTGGGGCCGACAGGAGTTGACATTTCTA	2132
Db	1740	CACAGATGCTGCCTGCCTGTGAGAGGGCGTGGGGCCGACAGGAGTTGACATTTCTA	1799
QY	2133	CCAGTCTCTACGAGAGAGCTGGGCCGCTGAGAGCTTTCACATCCCAAGAGAAG	2184
Db	1800	CCAGTCTCTACGAGAGAGCTGGGCCGCTGAGAGCTTTCACATCCCAAGAGAAG	1851

RESULT 3

```

Sequence 127829, Application US/09912293
GENERAL INFORMATION:
  APPLICANT: Rosen, et. al.
  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100
  FILE REFERENCE: PO-100
  CURRENT APPLICATION NUMBER: US/09/912,293
  CURRENT FILING DATE: 2001-07-26
  PRIOR APPLICATION NUMBER: 08/103,744
  PRIOR FILING DATE: 1993-08-09
  PRIOR APPLICATION NUMBER: 09/249,651
  PRIOR FILING DATE: 1999-02-12
  PRIOR APPLICATION NUMBER: 08/104,507
  PRIOR FILING DATE: 1993-08-09
  PRIOR APPLICATION NUMBER: 08/196,363
  PRIOR FILING DATE: 1994-02-15
  PRIOR APPLICATION NUMBER: 09/859,490
  PRIOR FILING DATE: 2001-05-18
  PRIOR APPLICATION NUMBER: 08/196,362
  PRIOR FILING DATE: 1994-02-15
  PRIOR APPLICATION NUMBER: 08/221,623
  PRIOR FILING DATE: 1994-03-31
  PRIOR APPLICATION NUMBER: 08/220,691
  PRIOR FILING DATE: 1994-03-31
  PRIOR APPLICATION NUMBER: 09/741,830
  PRIOR FILING DATE: 2000-12-22
  PRIOR APPLICATION NUMBER: 09/813,155
  PRIOR FILING DATE: 2001-03-21
  Remaining Prior Application data removed - See File Wrapper or PALM.
  SEQ ID NO 127829

```

```

: LENGTH: 352
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (351)..(351)
: OTHER INFORMATION: n is equal to a,t,g, or c
US-09-912-923-127829

```

Query Match	15.84:	Score 345.2:	DB 6:	Length 352:
Best Local Similarity	99.18:	Pred. No. 2.2e+68:		
Matches 347:	Conservative 0:	Mismatches 68:	Indels 0:	Gaps 0:
QY 1740	CAGTGTGCCTTCTTCACCGACTGGAGACACACATGTCATATACCCCGGTGCTCC	1799		

Db	1	CAGGTGACCTTTTCCCACTGACTGACGAGGAGCTGACATGTACATTAACGCCCGGGTGTGC	60
QY	1800	CTCCGAGAGGTCTGAGATATACCTGACCCCGCATGAGCGGGGAGACCTTCACGACAGCACT	1859
Db	61	CTCCGACAGAGGTCTGAGATATACCTGACCCCGCATGAGCGGGGAGACCTTCACGACAGCACT	120
QY	1860	GCAGCGCGGTGAAGATCCTCGCCGCTCAGAGACCTCATTTTGGTGTCCCGAGCGGGGTGAGA	1919
Db	121	GCAGCGCGGTGAAGATCCTCGCCGCTCAGAGACCTCATTTTGGTGTCCCGAGCGGGGTGAGA	180
QY	1920	TGTTATGCTATTGGCCTGAGAGAAGATTCTGAAGCCGAGCGGGGCCGAGTCATTGGCCT	1979
Db	181	TGTTATGCTATTGGCCTGAGAGAAGATTCTGAAGCCGAGCGGGGCCGAGTCATTGGCCT	240
QY	1980	CTTAAAAAGCCCGAGAGCTGACCTCCGATGAGGGGTGCTGTGTGATGCTGCCGTGTGGCAAA	2039
Db	241	CTTAAAAAGCCCGAGAGCTGACCTCCGATGAGGGGTGCTGTGTGATGCTGCCGTGTGGCAAA	300
QY	2040	GGAACATGTTGTGTGACCACTTTGAAAAATGAAAAACAGAGTGTGTCCTGG	2089
Db	301	GGAACATGTTGTGTGACCACTTTGAAAAATGAAAAACAGAGTGTGTCCTGG	350

RESULT 4

```

? Sequence 16014, Application US/10349781
? GENERAL INFORMATION:
? APPLICANT: Sellhammer, Jeffrey J.; Delisneane, Angelo M.
? APPLICANT: Stuart, Susan G.; Stuve, Laura L.
? APPLICANT: Mullany, Sara J.; Naughton, Rebecca E.
? TITLE OF INVENTION: POLYNUCLEOTIDES OF CELLS AND TISSUES OF THE FEMALE REPRODUCTIVE
? FILE REFERENCE: PD-1028-2 CON
? CURRENT APPLICATION NUMBER: US/10/349,781
? CURRENT FILING DATE: 2003-01-21
? PRIOR APPLICATION NUMBER: 09/540,764
? PRIOR FILING DATE: 2000-03-30
? PRIOR APPLICATION NUMBER: 08/992,868
? PRIOR FILING DATE: 1997-12-11
? PRIOR APPLICATION NUMBER: 60/032,838
? PRIOR FILING DATE: 1996-12-13
? PRIOR APPLICATION NUMBER: 08/734,050
? PRIOR FILING DATE: 1996-10-18
? PRIOR APPLICATION NUMBER: 60/006,111
? PRIOR FILING DATE: 1995-10-24
? PRIOR APPLICATION NUMBER: 08/733,814
? PRIOR FILING DATE: 1996-10-18
? PRIOR APPLICATION NUMBER: 60/005,864
? PRIOR FILING DATE: 1995-10-26
? PRIOR APPLICATION NUMBER: 08/763,920
? PRIOR FILING DATE: 1996-12-11
? PRIOR APPLICATION NUMBER: 60/008,794
? PRIOR FILING DATE: 1995-12-14
? PRIOR APPLICATION NUMBER: 08/772,783
? PRIOR FILING DATE: 1996-12-23
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 61458
? SOFTWARE: PERL Program
? SEQ ID NO 16014

```

```

: LENGTH: 288
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: misc.feature
:
: OTHER INFORMATION: Incyte ID No: hu00910748
:
: FEATURE:
:
: NAME/KEY: unsure
:
: LOCATION: 192, 263
:
: OTHER INFORMATION: a, t, c, g, or other
:
: US-10-349-781-16014

```

Query Match	13.1%	Score 286;	DB 9;	length 288;
Best Local Similarity	99.3%	Pred. NO. 5.3e-55;		
Matches 286; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

1. APPLICANT: INVENTOR: J. Craig et al.
2. TITLE OF INVENTION: COLLECTION OF CODING REGION SINGLE
3. TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISM (CSNPS) LOCATED ON EACH OF THE HUMAN
4. TITLE OF INVENTION: CHROMOSOMES, METHODS OF DETECTION, AND USES THEREOF
5. FILE REFERENCE: CL000896
6. CURRENT APPLICATION NUMBER: US/09/947, 907
7. CURRENT FILING DATE: 2003-03-17
8. PRIOR APPLICATION NUMBER: 60/241, 755
9. PRIOR FILING DATE: 2000-10-20


```

Db 125 TGGCAATGTAATCTGCTG-----GATGCTGGAGGACATCAATCACTGACTTGG 175
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 365 ACTACGGGATTTGAGGACGACTATTCATGATGAGGCGCCCTAGGCGGTGAGGCGACTCTG 424
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 176 GCATGTGTAAAGAGAAAGTCTTCCCGGAGACAAACCCGACCTTCTGCGNGAGCCCGG 235
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 425 GCTACGAGGCGCCGAGATATGAGGCTGCACTGATATGATGAGAGATGATGATGCTTCT 484
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 236 ACTACATAGTCCCGGAGATATGCTTACGACGCGCTATGAGGAGTGTGCTGATGATGCTG 295
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 485 CTTATGATGCTGCTTACGATGCTGCTGACGAGCAGCGCC 527
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 296 CTTTGGAGTTCTGCTGATGANAATGTTGGCAGACAGCTCC 338
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

RESULT 14

US-10-156-761-4356
Sequence 4356, Application US/10156761

```

GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIRAMA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHITAKI
APPLICANT: HATTORI, MASAHISA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 4356
LENGTH: 1497
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1497)
US-10-156-761-4356

```

Query Match 2.98; Score 64.2; DB 8; Length 1497;

Best Local Similarity 43.48; Pred. No. 0.00015;
Matches 294; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

```

Qy 23 CCGATGCCATAAAGATCTTCCGAGTTCGCGAGAGGCCAGCATGCTGCACGCGCTGC 82
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 164 CCGAGGAGCCCGTCTCCAAAGCCCGCTTACGCGCGAGGCCCACTCGGTGCGGCGCTCA 223
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 83 AGCAGCCCTGATCGTGGCGCTCATCGGATCAGCATCCAGCCGCTGCTGCGCCCTGG 142
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 224 ACCACATGCGGTGCTGCGCTGATGACACTCGCGGAGAGAGTGTGCGGAGCTCCACCG 283
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 143 AGCTGGCGCCCTCAGCAGCTCAACACCGTGTGTCGAGAGGCCAGAGATCTTCT 202
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 284 TCCGATCATGCTGATGAGTGTGTCGAGGCGCGACCATCCGGACTGTGCTCAACG 343
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 203 TTATACCCCTGGGACATGCTCAACCCCAAAATAGCTTACCAAGATCGCTCGGCGCTGG 262
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 344 CCGAGGCGCCCGGCGCCAGAGGCGTGTGATCATGCTCCGGGAGTGTGAGGCGCTGG 403
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 263 CCTACCTGACAGAAAGAAATCATCTTCTGAGCCTGAGTCCGAGCAACATCTGCTGT 322
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 404 CTTACTGACACAGACAGCAGCTGCTGACCGCGACATCAAGCCGCGCAAGCTCATATCA 463
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 323 GGTCCCTTGAAGTCAAGAGCATCAACATCAAGCTATCTGACTACGAGATTTGAGAGC 382
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 464 CGCAACAGGCTGCGGTGAAGGTCAATGAGACTTGGGATCGCGGCGCGCTGACGCGCGC 523
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

Qy 383 AGCATTCATGAGGGCGCCCTAGGCGTGGAGGACATCTGCTGACAGGCCAGAGA 442
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 524 AGTGCAGATGAGCAGACACCGCATGCTCATGGGACGCGCGCACTTCTGCGCGAGC 583
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 443 TCAGGCTCCGATGTATATGATGAGAGGTAGATATATGCTCTTATGAGATGCTCT 502
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 584 AGGGCTCGGAAAGGGCGGTGACACACCGCAGCGACTGTACGCGACGGTGTGCTGCT 643
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 503 ACAGATGCTGTACAGGACAGCGCCCTGCTACTGGGCGCCACACCACTCAATGCGCA 562
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 644 ACAGATCTCTGCGCTGCGCGCCCGCTTCAACCGGTGAGACCCGCTGTCCGTGTAC 703
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 563 AGCTGTCCAGAGGAGATCCCGCGCTTCTGGGCGACCGGAGAGTGCAGTCCGCGAGC 622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 704 AGCAGCTGACAGACATCCCGTGTCCCGCTCCGAGCTGCGGAGGAGGCGCGCGAGC 763
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 623 TGCAGGCGCTCATGATGAGTGTGAGACATGAGCCAGAGAAAGCAGCCGCTGCTGT 682
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 764 TCGAGGAGCTGCTGATGCTGCTGCTGCGCAAGATCCGAGACAGCGGTTCCAGACCGCG 823
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 683 CGGTGTGACCCAGATG 699
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 824 AGGAGATCGCGCGGCTG 840
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

RESULT 15

US-10-369-493-43167
Sequence 43167, Application US/10369493

```

GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 43167
LENGTH: 1069
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-10-369-493-43167

```

Query Match 2.88; Score 61; DB 8; Length 1069;

Best Local Similarity 45.98; Pred. No. 0.00072;
Matches 306; Conservative 0; Mismatches 340; Indels 21; Gaps 2;

```

Qy 28 GCCATGAAGAACTTCTCCGAGTTCGCGAGAGGCCAGCATGCTGCACGCGCTCACAC 87
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 121 GACAGGAGAGAGTGTGCGCTTCCGGCGGAGGCGTCCGCGGCTTCCGCTCAC 180
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 88 CCTGCTGCTGCGGCTCATGCGATCAGATCCACCGCGCTTGTGCTTGGCTGAGCTC 147
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 181 CAGAACATCTGTGACGCTCTACGATGTGAGAGAGGTGAGAGCGCTTCTCATGTGTG 240
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 148 GCGCGCTGACAGCAGCTCAACACCGTGTGCGAGAGAGCGCAGAGATTTCTTTATA 207
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 241 GAGCTGTGAGAGGCGCCACCTTCCACAGCTCATCAAGGAGGCG----- 285
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 208 CCCCCTGGAGACATGCTCAACCAAAATAGCTTACAGATGCGCTGCGCGCTGCTAC 267
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 286 CCGCTGCCCGGATGTACCGGGGCTCATCGCGCGCATCGCAGGCGCTGAGACAC 345
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 268 CTGACAGAGAAATCATCTTCTGTGACCTGAAGTGGAGCAACATCTGTGCTGTGCTC 327
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 346 GCGACATTCGCCACATCATCACCGGAGCTCAAGCCCGCCCAAGCTCATCACCAAG 405
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 328 CTTAGCGTCAAGAGACATCAATCAATCAATCTGACTACGAGGATTTGAGGAGAGTCA 387
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

QY 481 ThrTyrGlnLeuGlyAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
Db 1885 ACTTACACTGTGTGCTCCGGTACTTCTCGGGGTCCAGCCCTTCAGGACATGTTT 1944
QY 501 ProValArgProLeuAspThrGluProPheAlaSerHisThrAlaAsnProLysVal 520
Db 1945 CCGGTGCGGCGCTTGACAGGAAACCCCGGAGCAGCACAGGCGCAACCAAGGTG 2004
QY 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluLeuGlyThrGln 540
Db 2005 CCGAGGGGAGCTCCATCCGGAGCTGAGCATCATGATGAGAGAGTGGGACGACG 2064
QY 541 IleLeuIleHisGlnLeuSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 560
Db 2065 ATCTCATTCACACAGAAATCAGTCACTGATGCTCATGCTCATGCTCTCTCATCC 2124
QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSer 580
Db 2125 CCACCCCGGAGCTCCAGGTCCCTCAAGCTCCCGAGCTCCCGAGCAAGTCTTCC 2184
QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAla 600
Db 2185 AGGTGCTCTTCTCCACCCAGCTCCGAGACTCAGACATGCTATACATACCCGCTG 2244
QY 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHis 620
Db 2245 TCCGACAGGTCTGACATGATGATGACCCCATGAGCGGAGAGACCTTCAACCA 2304
QY 621 -GlnAlaValLys-IleLeuAlaValArgAspLeuIleTyrValProArgArgGly 640
Db 2305 GCAGGCGCGGAGAAATCTCCGCTCAGAGACTCATTTGGTCCCGAGGCGGTGAG 2364
QY 640 spValIleValIleGlyLeuGlnLysAspSerGluAlaGlnArgGlyArgValIle 660
Db 2365 ATGTTATTCGTCATGCGCTGGAGAGATCTGGCGCCGAGCGGCGGAGTCTATCCG 2424
QY 660 alleuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaValAla 680
Db 2425 CCTTAAAGCCGAGAGCTACCTCCGATGGGCGTGGTGGATGCTGCCCGGAGCA 2484
QY 680 YsaAspThrValLys-ThrPheGluAsnGlnAspThrGluThrPyrGluAlaVal 699
Db 2485 AGGACACTTGTGTGACCTTGAAGAAACACAGAGTGGTCCGCGCTGG 2544
QY 700 ArgGlyTyrGlyAlaArgGluPheAspIlePheTyrGlnSerTyrGluGlyArg 719
Db 2545 AGGGGCGGGGCGGCGAGGAGTTCGACATTTTTCACAGTCCATACGAGAGTGG 2604
QY 720 LeuGluAlaCysThrArgLysArg 727
Db 2605 CTGGAGGCTTGCACTCGCAAGAG 2628

```

RESULT 5

AAH77994
ID AAH77994 standard; DNA: 6045 BP.

AC AAH77994;

DT 13-NOV-2001 (first entry)

DE Nucleotide sequence of human protein kinase SGK258.

KM Human; protein kinase; cancer; immune disease; cardiovascular disease;

KM brain disease; neuronal disease; Alzheimer's disease; chromosome 15;

KM Parkinson's disease; multiple sclerosis; metabolic disorder;

KM peripheral nervous system disease; amyotrophic lateral sclerosis;

KM infection; ocular disease; migraine; pain; sexual dysfunction;

KM mood disorder; attention disorder; cognition disorder; hypotension;

KM hypertension; psychotic disorder; dyskinesia; transplant rejection; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..6045
FT /tag="a
FT /product="protein kinase"

PN W020016594-A2.

PD 13-SEP-2001.

PF 02-MAR-2001; 2001MO-US06838.

PR 06-MAR-2000; 2000US-0187150.

PR 29-MAR-2000; 2000US-0193404.

PR 13-NOV-2000; 2000US-0247013.

PA (SUGS-) SUGEN INC.

PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R.

DR WPI; 2001-536777/59.

PS P-PSDB; AAG67395.

XX Nucleic acids capable of encoding human polypeptides having a kinase or

PT kinase-like activity useful for diagnosing a disease selected from

PT cancers, cardiovascular disease and neuronal-associated diseases (e.g.

PT Alzheimer's disease).

PS Example 1; Fig 1E-F; 201pp; English.

XX The present sequence encodes a human protein kinase. The gene is
CC located on chromosome 15. The kinase polypeptides are useful for
CC diagnosing a disease or disorder selected from cancers (e.g. cancers
CC of tissues and cancers of hematopoietic origin), immune-related diseases
CC and disorders, cardiovascular disease, brain or neuronal-associated
CC diseases (e.g. Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis), metabolic disorders, peripheral nervous system diseases,
CC amyotrophic lateral sclerosis, viral infections, infections caused by
CC prions, infections caused by bacteria, infections caused by fungi,
CC ocular diseases, migraines, pain, sexual dysfunction, mood disorders,
CC attention disorders, cognition disorders, hypotension, hypertension,
CC psychotic disorders, dyskinesias, and organ transplant rejection.
CC Kinase inhibitors are useful for treating diseases and disorders
CC described above.

XX Sequence 6045 BP; 1359 A; 1779 C; 1707 G; 1200 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 1, 87e-275 Length: 6045

XX Score: 3579, 50. Matches: 688

XX Percent Similarity: 94.518 Conservative: 0

XX Best Local Similarity: 94.518 Mismatches: 1

XX Query Match: 93, 664 Indels: 39

XX DB: 22 Gaps: 1

US-09-836-392-21 (1-728) x AAH77994 (1-6045)

QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheSerGluPheArgGln 20

Db 3976 ATGCTGAGGACCTGCGGCGGCGAGATGCGATGAGAACTTCCAGATTCGCGGAGAG 4035

QY 21 AlSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40

Db 4036 GCCAGCATGCTGACGCGGCGGAGACACCTGATGCTGGCTCATCGGACATCAGCATC 4095

QY 41 HisProLeuGlyPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60

Db 4096 CACCGCTGCTTCCGCTCGGAGCTGCGGCGGCGGAGCTCAAGACCGTCTGCTCC 4155

QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnLysIleAla 80

Db 4156 GAGAGCGGAGAGATTTCTTATACCCCTGGAGACATGCTCAACCAAAATAGCC 4215

QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100

OY	721	GluHlaCysThrArgLyAsrGarg 728
DB	6245	GAGCGTTCACCTCCGAGAAGAGG 6268
 RESULT 7 AAK70889		
ID	AAK70889	standard; DNA; 18286 BP.
AC	AAK70889;	
XX		
DT	06-NOV-2001	(first entry)
XX		
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25701	
KW	Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer	
KM	Cytostatic; gene therapy; vaccine; metastasis; ds.	
XX		
OS	Homo sapiens.	
PN	WO200157182-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	17-JAN-2001; 2001WO-US01354.	
XX		
PR	31-JAN-2000; 2000US-0179065.	
PR	04-FEB-2000; 2000US-0180628.	
PR	24-FEB-2000; 2000US-0184664.	
PR	02-MAR-2000; 2000US-0186350.	
PR	16-MAR-2000; 2000US-0189874.	
PR	17-MAR-2000; 2000US-0190076.	
PR	18-APR-2000; 2000US-0198123.	
PR	19-MAY-2000; 2000US-0205515.	
PR	07-JUN-2000; 2000US-0209467.	
PR	28-JUN-2000; 2000US-0214886.	
PR	30-JUN-2000; 2000US-0215135.	
PR	07-JUL-2000; 2000US-0216647.	
PR	07-JUL-2000; 2000US-0216880.	
PR	11-JUL-2000; 2000US-0217487.	
PR	11-JUL-2000; 2000US-0217496.	
PR	14-JUL-2000; 2000US-0218290.	
PR	26-JUL-2000; 2000US-0220963.	
PR	26-JUL-2000; 2000US-0220964.	
PR	14-AUG-2000; 2000US-0224518.	
PR	14-AUG-2000; 2000US-0224519.	
PR	14-AUG-2000; 2000US-0224519.	
PR	14-AUG-2000; 2000US-0225213.	
PR	14-AUG-2000; 2000US-0225214.	
PR	14-AUG-2000; 2000US-0225266.	
PR	14-AUG-2000; 2000US-0225266.	
PR	14-AUG-2000; 2000US-0225267.	
PR	14-AUG-2000; 2000US-0225268.	
PR	14-AUG-2000; 2000US-0225270.	
PR	14-AUG-2000; 2000US-0225447.	
PR	14-AUG-2000; 2000US-0225757.	
PR	14-AUG-2000; 2000US-0225758.	
PR	14-AUG-2000; 2000US-0225759.	
PR	18-AUG-2000; 2000US-0226279.	
PR	22-AUG-2000; 2000US-0226681.	
PR	22-AUG-2000; 2000US-0226868.	
PR	23-AUG-2000; 2000US-0227182.	
PR	23-AUG-2000; 2000US-0227009.	
PR	30-AUG-2000; 2000US-0228924.	
PR	01-SEP-2000; 2000US-0229287.	
PR	01-SEP-2000; 2000US-0229343.	
PR	01-SEP-2000; 2000US-0229344.	
PR	01-SEP-2000; 2000US-0229345.	
PR	05-SEP-2000; 2000US-0229509.	
PR	05-SEP-2000; 2000US-0229513.	
PR	06-SEP-2000; 2000US-0230437.	
PR	06-SEP-2000; 2000US-0230438.	
PR	08-SEP-2000; 2000US-0231242.	
PR	08-SEP-2000; 2000US-0231243.	
PR	08-SEP-2000; 2000US-0231244.	
PR	08-SEP-2000; 2000US-0231413.	

[illegible]

PR	17-NOV-2000;	2000US-0249329.
PR	17-NOV-2000;	2000US-0249330.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251997.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HOMA-) HUMAN GENOME SCI INC.	
XX		
PL	Rosen CA, Barash SC, Ruben SM;	
DR	WPI: 2001-483426/52.	
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	
PS	Disclosure; SEQ ID NO 25701; 3071pp + Sequence Listing; English.	
XX		
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)	
CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/hematopoietic-related diseases, especially	
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703	
CC	to AAK87694 represent human immune/hematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK82169	
CC	represent sequences used in the exemplification of the present invention.	
SQ	Sequence 18286 BP; 4372 A; 4856 C; 4941 G; 4117 T; 0 other;	
Alignment Scores:		
Pred. NO.:	1,27e-113	Length: 18286
Score:	1560.50	Matches: 319
Percent Similarity:	68.15%	Conservative: 2
Best Local Similarity:	67.73%	Mismatches: 4
Query Match:	40.83%	Indels: 147
DB:	22	Gaps: 2
US-09-836-392-21 (1-728) x AAK70889 (1-18286)		
OY	346 ValProVal---IlleLysLysAsnSerTYrLeuValLeuAlaGlyLeuAlaAspLysLeu 364	
Db	13505 ATCCCTGTCTCCTTCCTTCAAGTAATTCCTACCTGGTTTGCGGGCTCCCATGGCGTT 13564	
OY	365 ValAlaValPheProValValAArgGLyThrProLysASerSerCySerTYrLeuGlySer 384	
Db	13565 GTGGCTGTGTTCCCGGGGGGGGCCAACCCCAAAGAAGCGTGCCTCCTACTGTCCTGA 13624	
OY	385 HisThrLaasnaArSerLysPheSerIleAlaAspGLuaSPALaArgLnAsnProTYr 404	
Db	13625 CACAAGACCAACAGGTCTCAAGTTCAGATCGCGGATGAAGACGACGAGCAACCCTCAC 13684	
OY	405 ProValLysalaMetGlualValaAsnserGlySerGeluaValITrptySerasnGLyPro 424	
Db	13685 CCACGTGAAGGCAACGAGGAGGTCACACAGCGGCTCTGAGGTGTGGTAGCAAGCATGGCGCG 13744	

PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0234001.
 PR 14-SEP-2000; 2000US-0234010.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249300.
 PR 17-NOV-2000; 2000US-0249309.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 FI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 38916; 3071pp + Sequence listing; English.
 XX
 CC AAK64951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK67694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
 CC represent sequences used in the exemplification of the present invention.
 CC
 XX
 SQ Sequence 18286 BP; 4372 A; 4856 C; 4941 G; 4117 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.27e-113 Length: 18286
 Score: 1560.50 Matches: 319
 Percent Similarity: 68.15% Conservative: 2
 Best Local Similarity: 67.73% Mismatches: 4
 Query Match: 40.83% Indels: 147
 Gaps: 2
 DB:
 US-09-836-392-21 (1-728) x AAK84104 (1-18286)
 QY 346 ValProval---IleIysIysAsnSerIyrluValIleuAlAGlyleuAlaaspGlyIleu 364
 DB 13505 ATCCCTGTCCTCTCTGATTCAGATTCCTACCTGCTTAGCGGGCTCCGACATGGGCTT 13564
 QY 365 ValAlaValAlaPheProValValArgGlyThrPolysAspSerCySerIyrluCySer 384
 DB 13565 GGGGCTGTGTTCCCGTGTGGGGGACCCCAAGAGACAGTCTCCTAGCTGTGCTCA 13624
 QY 385 HistHraAlaAsnArgSerIyrluAlaaspGluAspAlaArgGlnAsnProTyr 404
 DB 13625 CACACAGCCACAGCGTCCAAAGTTCAGCATCCGGATGAAGACGACAGGACAGACCCCTAC 13684

[illegible]

Db	13626	CACAGACCCACAGGCTCCAGATTACGACTCGGGATGAAAGCGCAGCGGACAGAACCCCTTAC	13685
Qy	405	ProValIySaIaMeGluValValaIaSaSerGIySerGIuValITPrTySerAsnGIyPro	424
Db	13686	CCAGTGAAGGCCATGTGAGAGTGTGTCAACAGCGGCTGTGAGGTCTGGTACAGCAATGGGCG	13745
Qy	425	GIyLeuLeuValIlaSPcysAlaSerLeuGIuIleCySaIaGlyLeuGIuProTyzMet	444
Db	13746	GGCCCTCTTTCATGACATGTGTCCCTCCCTGGAGATCGACAGCGGCTGGAGCCCTACATG	13805
Qy	445	AlaProSerMetValThSerValValySerSerGIuIyArgGIyGIuGIuValVal	464
Db	13806	GGCCCTCCAGGTTCAGTCACATCTGTGTGACACTGTGTGAGGCGAGGGGAGAGGTGTCTC	13865
Qy	465	TrpCysLeuAspAspLySaIaIaSaSerLeuValImeCTyHISerThThTyGIuLeu	484
Db	13866	TGGTCCCTGTGAAGACAAGGCCAACTCTTGGTGAATGTACACTCCACCACTACAGAGCTG	13925
Qy	485	CysAlaArgTyRPhcCysGIyValProSerProLeuArgAspMetPheProValArgPro	504
Db	13926	TGTGCCGGTACTTCTGGGGGGTCCCAAGCCCTCCAGGAGACATGTTCCTCCGTGGGGCC	13985
Qy	505	LeuAspThrGIuProProAlaIaSerHISThrAlaSaProLyValProGIuIyAsp	524
Db	13986	TTCGACAGGAAACCCCGGACCCAGCCACACAGGCCAAACCCMAAGTGTCTAGGGGGGAC	14045
Qy	525	SerIleAlaAspValSerIleMetTySerGIuIyLeuGIuIleGIuIleLeuIleHIS	544
Db	14046	TTCATCGGGGAGTGAAGCATCTGTACAGTGTGAAGTGGAGCTGGGACGACAGATCTGTACAC	14105
Qy	545	GIuGIuSerLeuThrAspTyCysSerMetSerSerTySerSerSerProArgIuGIu	564
Db	14106	CAGGATCACTACAGTACAGTACAGTCCCAATGTCTCTTACTCTCATCCCAACCCGGCCAG	14165
Qy	565	AlaAlaArgSerProSerSerLeuProSerProAlaSerSerSerSerValProPhe	584
Db	14166	GCTGCAGGTCCCTCCCTCAAGCCCTCCCAAGCTCCCAAGAACTTCTCCAGTGTGCTTTC	14225
Qy	585	SerThrAspCysGIuAspSerAspMetLeuHISThrProGIuAlaIaSerAspArgSer	604
Db	14226	TCCACCGAGTGCAGGACTCAACAATGTACTACAGCGCGGTGTGCTCCCAAGGTCT	14285
Qy	605	GIuHISAspLeuThrProMetAspGIyGIuThrPheSerGIuHISLeuGIuAlaValIyS	624
Db	14286	GAGCATGACCTTACCCCTTGAAGCGGGAGACTTTCAGCCACACCTGAGGCGCGTGAAG	14345
Qy	625	IleLeuAlaValaIArgAspLeuIleTrpValPro-----	635
Db	14346	ATTCCTGCGCGTGAAGACTCTCATTTGGGG-CCCAAGGTACGTTCCCGAGGTGAAGGCAC	14404
Qy	635	-----	635
Db	14405	CATCCAGGCGACGCCCATGTCTGTCTGTGGGACAGACAAAGGGAAGCCCTGTGCT	14464
Qy	635	-----	635
Db	14465	TCCCTACAGGTATGTGGGAGAGCAGTGTCTTCTTCCAAAGAGGTGTCTTCTCCAT	14524
Qy	635	-----	635
Db	14525	TGGGATGGAAGGATCGCTTTAGCAAAATCTAAAGATGGCGCGACAAAAGCTTATTGCT	14584
Qy	635	-----	635
Db	14585	GCCAGAGACTTAAAGTGCACACAAGAAAGCACAGTCAACTGTGACTCAGTTTCCAC	14644
Qy	635	-----	635
Db	14645	TGAGAAATTTCTCTGTGCTTGAAGTCTGTGCCCTCCCACTACTTAATTCCTTTGACAG	14704

[illegible]

Db 7149 CTTGTGACATTAAGATTGCGACTACGAAATGACCGCAACTGCTCAAGTGGACCT 7208
 QY 135 LeuGlyValGluGlyThrProGlyTyrGlnAlaProGluLeu-----ArgProArg 151
 Db 7209 AAGAGATTGGTGGCCGAGGTTTATGCTCCGGAGATCAATACGCTACAAATGGTGAG 7268
 QY 152 IleValTyrAspGluLys----- 157
 Db 7269 GAGGAATTAAGTGAAGGATTTATTTATTTACTTCTCATTAATCATATGCTTATC 7328
 QY 158 -----ValAspMetSerTyrGlyMetValLeuTyrGluLeuLeuSer 172
 Db 7329 GATCTCTTGTGACGAGTGGAGCTGCTCTCATTTGGAATGTCAATATAGAAATATACGT 7388
 QY 173 GlyGlnArgProAlaLeuGlyHisHisGlnLeuGlnIleAlaLysLysLeuSerLysGly 192
 Db 7389 TTACGACAACTTCGAGAGGCCAGAG-----TCTATTAAGAGTGCATCTTGGAGGGT 7442
 QY 193 IleArgProValLeuGlyGlnProGluGlnValGlnPheArgArg-----LeuGlnAlaLeu 211
 Db 7443 AGTGGCCCGCTTGACCTCAA---AGGGAACCCAGTTTCCCACTGCTGTTGGATCTT 7499
 QY 212 MetMetGluCysTyrPaspThrLysProGluLysArgProLeuAlaLeuSerValLysSer 231
 Db 7500 ATGCTCTGTGTGGACGACAGACAGCTCTGCGACAGCCGACAGTCAAGATTTGTTCC 7559
 QY 232 GlnMetLysAspProThrPheAlaThrPheMet-----Tyr 243
 Db 7560 ATACTAAGTGCACCGAGTGCATCCACTCCTCGATGCTGTGCCANTGCCGATAGCGAG 7619
 QY 244 GluLeuCysGlyLys-----GlnThrAlaPhePheSerSerGlnGlyGlnIleTyrThr 262
 Db 7620 AAGATGTCTGTGGGCTTTTCAAGTCTGTGTGTATGGGCGATGATGAAGATGTGTGT 7679
 QY 263 ValValPheTyrPaspGlyLysGlnGlnSerArgAsnTyrThrValValAsnThrGluLys 282
 Db 7680 CTGTAAGCTGTGGTACCTCTCTCGGCTCTCGCATTTGATTCATTCAGTCTCAACCTCG 7739
 QY 283 Gly-----LeuMetGluValGlnAlaArgMetCys----- 292
 Db 7740 GGCAGGCTACGACAGTGCACAGCATGAGTTCTCCGACAGCCAGGTTGCTCCGCC 7799
 QY 293 -----ProGlyMet 295
 Db 7800 AAGACACCCGAAACGATGCAATTCACAGTCCGCTGCTGCCCAACGTTTGCCCAAGATG 7859
 QY 296 LysValSerCysGlnLeuGlnValGlnArgSerLeuTyrThrAlaThrGlnAspGlnLys 315
 Db 7860 AAGATGTCTGTGGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7919
 QY 316 IleTyrIleTyr----- 319
 Db 7920 CTGCACGCTATAGACATCCACCTATGCCCACCTGTTTCTTACATGCTCATCCCAAC 7979
 QY 319 ----- 319
 Db 7960 ATTAAGTCAGCTGTGATAGCTACTACTACATGAGAAAATAGCTCGGCTGCTGCGA 8039
 QY 320 ThrLeuLysGlyMet-CysProLeuAsnThrProGlnIleAlaLeuAspThrProAla-- 338
 Db 8040 ACGCATACGATGCGGTCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8090
 QY 339 -----ValValThrCysPhe 343
 Db 8091 TGCCGATTGGCCGAGGCTCTTTTGTGCTTACAGAGATCTGTCTGTGTGTGTGTGTGT 8150
 QY 343 eleuAlaValProVal-----IleLysLys 351
 Db 8151 GCCGCTTGTCTCGT 8210
 QY 351 sasnSerTyrLeuValLeu-----AlaGlyLeuAlaAspGlyLe 364

Db 8211 AATCTATATTGCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8270
 QY 364 uValAlaValPheProVal-----ValArgGlyThrProLysAspSerCysSe 380
 Db 8271 GATTAATGTTTTCCTGTTGAACGAGATGAGTTTCCGACAT-----CA 8315
 QY 380 rTyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaAr 400
 Db 8316 GCGTTTATGC---CACAGCGAGAACCAATCTCATC-----GAGGAC----- 8355
 QY 400 gGlnAsnProTyrProValLysAlaMetGluValValAsnSerLysSerGluValTPTy 420
 Db 8356 -----GTCAGAGTGGCTCGAATGTGCAATATGAAACCATGTGTTCAG 8399
 QY 420 rSerAsnGlyProGlyLeuLeuVal----- 428
 Db 8400 TTGCTATATCCCGGT 8459
 QY 429 -----IleAspCysAlaSerLeuGluIleCys-----ArgArgLeuGluProTyrMetAlaPr 446
 Db 8460 TAAGCTGAGCTGCTCAAAAGCTGCTGCCCTGCTCCGAGTCTCGCAAGATATTGCGATGA 8519
 QY 446 oSerMetValThrSerValValCysSerSerGluGlyArgGlyGluValValTPTCy 466
 Db 8520 TGACAGTGTGAACCTTATTAATGC----- 8544
 QY 466 sleuAspAspLysAlaAsnSerLeuValMetTyrHisSerThrTyrGlnLeuCysAl 486
 Db 8545 -----CAGATTGTGGACTGGCGGCTCACATTCAGAACGTAT---ATTGGAC 8591
 QY 486 aArgTyrPheCysGlyValProSerProLeuArgAspMetPheProVal-----Ar 503
 Db 8592 GACGTGGGCTGCTTAATGTGGCGGAGACTGCACACCTGGCCCATAGATGTTTCCG 8651
 QY 503 gProLeuAspThrGluProProAlaAlaSer----- 513
 Db 8652 ACCCTATGAATGAGTGAAGTTGGCGAATTTGCGATTATGTGCTGATGCTGATT 8711
 QY 514 -----HisThrAlaAsnPr 518
 Db 8712 CTATTCCTTACAGATCAATCTATTAATACCTTTCTTAAGACATATTCGCTGATCG 8771
 QY 518 oLysValProGluGluAspSerIleAlaAspValSerIleMetTyrSerGlnGluLeuG 538
 Db 8772 CCAGATGGAAGACGATATGCTCTGTGATCTGTGATTA-----TGTGG 8816
 QY 538 yThrGlnIleLeuIleHisGlnGlu 546
 Db 8817 ACTCCGACAGTCAATCCACCAAGAG 8841
 RESULT 13
 ABL09495
 ID ABL09495 standard; cDNA; 6927 BP. ---
 AC ABL09495;
 XX
 DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster expressed polynucleotide seq ID NO 22967.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 KW Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.

QY 429 -IleAspCysAlaSerLeuGluIleCys---ArgArgLeuGluProTyrMetAlaProSe 447
 Db 6501 GCTGGAGTGGCTCAAGAGCTGCGCCCTGCGAGATCTGCAAGATATGGATTGATGAGA 6560
 QY 447 mCetValThrSerValValCysSerSerGluGlyArgGlyGluValValITPCCysLe 467
 Db 6561 GCATGTGAACTTATTAATAATC----- 6582
 QY 467 uAspAspLysAlaAsnSerLeuValMetTyrHisSerThrThrTyrGluLeuCysAlaAr 487
 Db 6583 -----CAGATTTCGGAGCTGGCGGCTCACAAATTCAGACTGAT---ATTGGAGCAC 6632
 QY 487 gTyrPheCysGlyValProSerProLeuArgAspMetPheProVal-----ArgPr 504
 Db 6633 CTGGGGCGCTTAATTTGGTGGCCGAGATCCACACCTGGCCCATTTAGTTTCCGAGC 6692
 QY 504 cLeuAspThrLupProProAlaAlaSerHisThrAlaAsnProLysValProGluGlyAs 524
 Db 6693 CTATGAAATATGAAATCAATATATATTAACCTTTCTTAAGACAAATGTTCCG----- 6744
 QY 524 pSerIleAlaAspAlaSerIleMetTyrSerGluGluLeuGlyThrGlnIleLeuIleH1 544
 Db 6745 -CTGATCCGCCAGATCGGAGAACAGATAT---CGCTCTGTGATCTCTCATATGTGGACTC 6800
 QY 544 sGlnGluSerLeuThrAspTyrCysSerMetSerTyrSerSerSerProProArgG1 564
 Db 6801 CGCAGAGCATCCAC-----AAGAGCTCTGCGCTCGACACACCAACCCACATGG 6848
 QY 564 nAlaAlaArgSer 568
 Db 6849 TGCAGCCAAATCC 6861
 RESULT 14
 ID AAD26452 standard; cDNA; 2693 BP.
 XX AAD26452;
 AC AAD26452;
 XX 26-MAR-2002 (first entry)
 DE Human kinase PKIN-5 cDNA.
 XX
 KW Human; kinase; PKIN-5; cancer; leukemia; adenocarcinoma; osteoarthritis;
 KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
 KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
 KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
 KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoporosis;
 KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
 KW rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
 KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
 KW cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
 KW congestive heart failure; ischaemic heart disease; lung tumour; gout;
 KW fatty liver; Niemann-Pick's disease; gene therapy; ss.
 KW
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 66..2036
 FT /tag= a
 FT /product= "Human PKIN-5 protein"
 XX
 PN MO200196547-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 14-JUN-2001; 2001WO-US19444.
 XX
 PR 15-JUN-2000; 2000US-212073P.
 PR 23-JUN-2000; 2000US-213467P.
 PR 30-JUN-2000; 2000US-213651P.
 PR 07-JUL-2000; 2000US-218372P.
 PR 13-JUL-2000; 2000US-218372P.
 PR 25-AUG-2000; 2000US-228056P.

XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Lal P, Bandman O, Borowsky M, Au-Young J, Lu Y,
 PI Gandhi AR, Tribouley CM, Walla NK, Yao MG, Lu DM, Greenwald SR,
 PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang Y,
 PI Baughn MR, He A, Thornton M, Hatfield A, Patterson C, Gururajan R,
 PI Lo TP, Khan F, Reardon SA, Azimzai Y, Policky JL, Ding L,
 PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH,
 XX
 DR WPI, 2002-090207/12.
 DR P-PSDB; AAE16359.
 XX
 PT New polypeptides, useful for diagnosing, treating or preventing
 PT disorders of growth and development, cardiovascular and lipid, and
 PT diseases such as cancer, comprise human kinase polypeptides
 XX
 PS Claim 5, Page 178-179; 197pp; English.
 XX
 CC The invention relates to human kinase PKIN proteins and their
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for
 CC treating a disease or condition associated with decreased expression of
 CC PKIN and a composition comprising PKIN antagonist is useful for treating
 CC a disease or condition associated with overexpression of PKIN. The
 CC disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma,
 CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease), immune disorder
 CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
 CC atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
 CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
 CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
 CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
 CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts), cardio
 CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
 CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
 CC ischaemic heart disease, chronic bronchitis, lung tumours), lipid
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
 CC hypocholesterolemia, obesity). PKIN DNA is useful for assessing
 CC toxicity of a test compound and in gene therapy. The present sequence
 CC is human PKIN-5 cDNA.
 XX
 SO Sequence 2693 BP; 857 A; 475 C; 550 G; 811 T; 0 other;
 XX
 Alignment Scores:
 Pred. NO: 1.53e-16 Length: 2693
 Score: 330.50 Matches: 108
 Percent Similarity: 47.34% Conservative: 79
 Best Local Similarity: 27.34% Mismatches: 150
 Query Match: 8.65% Indels: 59
 DB: 24 Gaps: 14
 US-09-836-392-21 (1-728) x AAD26452 (1-2693)
 QY 10 AlameLysAsnPheserGlu-----PheArgGlnGluAlaSerMet 23
 Db 162 GCTGTGAAGATTTTAAATAAACATACATCAGCTGTTAAGCAAGAGCTTGCTG 221
 QY 24 LeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIleHisProLeu 43
 Db 222 CTTTGCCACCTCCACACCCCGAGTTGATATCTTGGCGAGCTGGATTCGCCGG 281
 QY 44 CysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSerGluAsnAla 63
 Db 282 ATGTTGTGATGAGTACCTCCCAAGAGTTCTTGATATCTTGGCGAGCTGGATTCGCCGG 341
 QY 64 ArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnLysIleAlaTyrGlnIle 83
 Db 342 GCCAGC-----CTCAGTAAAGCCCTTAACAGACAGATTCACATCCACGTA 386
 QY 84 AlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeuLysSerAsp 103


```
QY 84 AlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeuLysSerasp 103
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1480 GCTGATGTTGAGATGACTCTCCACTCAGCCATGATTATATACCGAGCCTGAACCCAC 1539
QY 104 AsnIleLeuValTrpSerLeuAspValLysGluHisIleAsnIleLysLeuSerAspTyr 123
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1540 AATGGCGCTTTCACACTGTATCCCAATGCTGCATCATTTGCAAGATGCTGACTAC 1599
QY 124 GlyIleSerArgGlnSerPheHisGluGlyAlaLeuGlyValGluGlyThrProGlyTyr 143
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1600 GGCATGCTCAGTCTGCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1659
QY 144 GlnIleProGluIle---ArgProArgIleValTyrAspGluLysValAspMetPheSer 162
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1660 CGTGCACCTGAGTGGCCAGAGAAATGCTATTATTAACCAACAGCTGATGTTATTCOA 1719
QY 163 TyrGlyMetValLeuTyrGluLeuLeu---SerGlyGlnArgProAlaLeuGly----- 179
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1720 TTGGTTACTACTCTATGACATTTTGCACACTGAGAGTAGAATAGTAGAGGTTGAAG 1779
QY 180 -----HisHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgPro 195
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1780 TTTCCAATGAGTTGATGAAATAGAAATACAGAGAAATTAAGTATCACTGATCAAGTAAAGAA 1839
QY 196 ValLeuGlyGlnProGluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGluCys 215
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1840 TATGGTTGTGCCCA-----TGCCCTATGTTGAGAAATTAATTAACAGAGT 1887
QY 216 TrpAspThrLysProGluLysArgProLeuAlaLeuSerValIleSerGlnMetLysasp 235
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1888 TTGAAGAAGAAATCCCAAGAAAGAGCTCTGCTGAGCTTGTGACATTTGAATCA 1947
QY 236 ProThrPheAlaThrPheMetTyrGluLeuCysGlyLysGlnThrAlaPhe----- 253
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1948 GCT-----GAATAGCTGTCTCT-GACAGAGACGCAATTTATTTACC 1985
QY 254 -----PheSerSerGlnGlyGlnGlu-----Tyr 261
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1986 TAAAAACGTAAATTGTTGAATGCAATGCTGCTACACATCAACAGAGAAATGCAAGCAT 2045
QY 262 ThrValValPheTrpAspGlyLysGluGlnLysSerArg-AsnTyrThrValIleAsnThrG1 281
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2046 TTGGCTGGGCTGTGGCACACGACAGACAGACAGCTCTCATTTCTTGACTTAATACTGA 2105
QY 281 uLysGlyLeuMetGluVal-----GlnArgMetCysCysProGlyMetLysValSe 298
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2106 AGGATACACTTCTGAGAGAGTTGCTGATAGATGATGATGATGATGATGATGATGATGAT 2157
QY 298 rCysGlnLeuGlnValGlnArgSerLeuTrp-----ThrAlaThrGluAspGlnLysI1 316
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2158 -GTGCATCTCTCTGTGAAGAGAAAGCTGATTTGCTGTGGACACAGCTCTGACTCTT 2216
QY 316 eTyrIleTyrThrLeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspPh 336
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2217 COTGGTCAATCAATACGAAAGATGGGAAAAGAGACATACCCCAAGAAAGATGATGATTC 2276
QY 336 rProAlaValIleThrCysPheLeuAlaValProValIleLys-----LysAsnSerTy 354
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2277 r-----GTCACTTGTGTTGATGCAATTCCTTTTCCAAGCAAGCAAAACAAAAA 2327
QY 354 rLeuValLeuAlaGlyLeuAlaAspGlyLeuValAlaValPhe 368
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2328 TTTTCTTTGCTTGGAAACCGCTGATGCAAGTATGCAATTTT 2370
```

Search completed: April 15, 2003, 07:32:01
Job time : 365 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - nucleic search, using frame_plus.p2n model

Run on: April 15, 2003, 07:22:23 ; Search time 1722 Seconds

(Without alignments)
6846.872 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822

Sequence: 1 MLRLRLRLNDANKNSERFQ.....IFYSYERLGLACTRRKR 728

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-O=/cgn2.1/USPTO_SPOOL/US09836332/runat_08042003_090540_23413/app-query.fasta.1.903
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=humand0.cdi -LIST=45
-LOCALALIGN=200 -NHR SCORE=pcr -THR MAX=100 -THR MIN=0 -MODE=LOCAL
-OFFPRT=pts -NHR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09836332 -ECGN=1.1.1426 -etunat_08042003_090540_23413 -NCPu=6 -ICPU=3
-NO_XUPX -NO_WMAP -LANG=QUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMECUT=120
-MARL.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-GAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_estbun:*
3: em_estcin:*
4: em_estcnu:*
5: em_estcov:*
6: em_estcpl:*
7: em_esttro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_ham:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3578.5	93.6	3509	11 BC020379	BC020379 Homo sapi
2	1183.5	31.0	863	13 B1100538	B1100538 Homo sapi
3	1035	27.1	611	13 B1758444	B1758444 Homo sapi
4	1015	26.6	638	10 BB613947	BB613947 Homo sapi
5	948.5	24.8	1066	12 BF796030	BF796030 Homo sapi
6	943	24.7	1835	11 BC018934	BC018934 Homo sapi
7	907	23.7	562	12 BE899922	BE899922 Homo sapi
8	893.5	23.4	631	13 B1042551	B1042551 Homo sapi
9	883.5	23.1	737	12 BF301427	BF301427 Homo sapi
10	879.5	23.0	939	12 BF301427	BF301427 Homo sapi
11	878	23.0	864	12 BF583283	BF583283 Homo sapi
12	871	22.8	715	10 BB649591	BB649591 Homo sapi
13	836	21.9	996	13 B1468840	B1468840 Homo sapi
14	790	20.7	707	17 A2358773	A2358773 Homo sapi
15	786	20.6	484	14 BQ335590	BQ335590 Homo sapi
16	774	20.3	464	12 BG146794	BG146794 Homo sapi
17	767	20.1	461	12 BE234123	BE234123 Homo sapi
18	741	19.4	465	10 AM502376	AM502376 Homo sapi
19	721	18.9	1072	14 BQ064084	BQ064084 Homo sapi
20	717	18.8	471	10 AM500039	AM500039 Homo sapi
21	680	17.8	572	13 B1344463	B1344463 Homo sapi
22	678	17.7	481	12 BG061416	BG061416 Homo sapi
23	676	17.7	1091	12 BG388614	BG388614 Homo sapi
24	621	16.2	419	13 B1965906	B1965906 Homo sapi
25	582	15.2	358	14 AM505103	AM505103 Homo sapi
26	575.5	15.1	625	10 BQ386806	BQ386806 Homo sapi
27	569.5	14.9	673	12 BF303665	BF303665 Homo sapi
28	569	14.9	370	13 B1964647	B1964647 Homo sapi
29	511.5	13.4	464	14 BQ336146	BQ336146 Homo sapi
30	495	13.0	515	17 B15482	B15482 Homo sapi
31	493	12.9	728	12 BE795256	BE795256 Homo sapi
32	473.5	12.4	415	9 AA103505	AA103505 Homo sapi
33	471	12.3	375	9 AA879924	AA879924 Homo sapi
34	470	12.3	557	12 BF813830	BF813830 Homo sapi
35	455.5	11.9	343	13 B1025188	B1025188 Homo sapi
36	445.5	11.7	625	10 BB660785	BB660785 Homo sapi
37	428	11.2	473	10 AM502395	AM502395 Homo sapi
38	400	10.5	565	9 AA667408	AA667408 Homo sapi
39	392	10.3	860	10 BE544646	BE544646 Homo sapi
40	386	10.1	660	10 BB205937	BB205937 Homo sapi
41	349	9.1	233	10 AM504559	AM504559 Homo sapi
42	322	8.4	233	10 BE860033	BE860033 Homo sapi
43	320	8.4	363	12 BE860033	BE860033 Homo sapi
44	317	8.3	534	13 BM363561	BM363561 Homo sapi
45	306	8.0	467	17 A2771756	A2771756 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS BC020379 3509 bp mRNA
DEFINITION Homo sapiens, clone IMAGE:3463102, mRNA.
ACCESSION BC020379
VERSION BC020379.1 GI:18042971
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 3509)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian

OY	378	ercYsserTylencYsserHisTrrAlaAsnArgSerLysHsheRilLeaAlaapGlu	398
Db	661	gCTeCTCCTACCTGTCGCTCCACACAGC - AACAGTCCAGATTTCATCCC - GACGAG	718
OY	398	spAlaargLInasnpProTyrProValLysAlaMetGluValAlaAsnSerGlySerGluV	418
Db	719	ATGCACGGAG - AACCTTACCCAGTGAAGGCAATGGCAGTGGTCAAC - AGTGGCTGTGAGG	776
OY	418	aLTpTYrSerAsnGlyProGlyLeuValIlaAspCysAlaSerLeuGluIleCysA	438
Db	777	TCTGGTACGACACAGGGCT - GGGCTCCCTGCATCATGATGCCGACATCGTGCACATATACGA	835
OY	438	rGArgLeuGlu	441
Db	836	GCGCGATGAC	846

LOCUS	BT158444	611 bp	mRNA	linear	EST 25-SEP-2001
DEFINITION	603022572F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5193448 5', mRNA sequence.				
ACCESSION	BT158444				
VERSION	BT158444.1	GI:15750022			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 611)				
TITLE	NIH-MGC http://mgc.ncl.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				

```

FEATURES
source
    Tissue Procurement: Life Technologies, Inc.
    cDNA Library Preparation: Life Technologies, Inc.
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LMU at:
    http://image.llnl.gov
    Plate: LMU1483 row: 0. column: 17
    High quality sequence stop: 608.
    Location/Qualifiers
        1. 611

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5193448"
/clone_id="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-et
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
111 a 213 c 169 g 118 t

```

Alignment Scores:	
Pred. No.:	2,66e-93
Score:	1035.00
Percent Similarity:	98.52%
Best Local Similarity:	98.03%
Query Match:	27.08%
DB:	13
Length:	611
Matches:	199
Conservative:	1
Mismatches:	2
Indels:	1
Gaps:	0

US-09-836-392-21 (1-728) x B1758444 (1-611)

OY	408	AlmetgluValAlaAsnSerGlySerIuValTPrTySerAsnGlyProGluLeu	427
Db	4	GCCATGGAGGTGGTCAACAGCGGCTCTAAGTCTGTGTACAGCATGGGCGGGCTCTT	63
OY	428	ValIleAspCysAlaSerLeuGluIleCysArgArgLeuGluProTyrMetAlaProSer	447
Db	64	GTCATCGACGTGGCTCCCTGGAGATGTGCAMGGCGGTGAGACCCTACATGGCCCCCTTC	123
OY	448	MetValThrSerValValCysSerSerGluGlyArgGlyGluValValTProGlyLeu	467
Db	124	ATGGTTAGGTAGTGTGTGTGCAGCTTGAAGGCAAGGGAGAGAGTGTGGTGGCTG	183
OY	468	AspAspIysAlaAsnSerIleuValMetTyrHisSerThrThrTyrGlnLeuCysAlaArg	487
Db	184	GATGCAAGGCCCACTCTTGGTATGTATGCACATCCACCTACACACTGTGTGCCCCG	243
OY	488	TyrPheCysGlyValProSerProIleuArgAspMetPheProValArgProIleuAspThr	507
Db	244	TACTCTCTCGGGGTCGCCAGCCCCCTCAGGGCAAGTTCCTCCGCGCTTGGACAG	303
OY	508	GluProProAlaIleSerHisThrIleAsnProIysValProGluGlyAspSerIleAla	527
Db	304	GAACCCCGGGCAGCCAGGCCACAGGCCAACCAGAAAGTGCTGAGGGGGAATCCATCCG	363
OY	528	AspValSerIleMetTyrSerGluIleuGluGlyThrGlnIleLeuIleHisGlnIleSer	547
Db	364	GACGGACATCATGTACAGTAGAGAGCTGGGACAGCAGATCTGATCCACACAGAAATCA	423
OY	548	LeuThrAspTyrCysSerMetSerSerTyrIleSerSerProProArgGlnAlaIleArg	567
Db	424	CTCATGTACTGTGTCCATGTGCTCTCTACTCTCTATGCCACCCCGGACGGCTCCAGG	483
OY	568	SerProSerSerIleuProSerSerProAlaSerSerSerValProPheSerThrAs	587
Db	484	TGGCCCCGTCAGAGCCGCCAGAGTCCCGACAGAAAGTCTTGCAGATGTGCCTTTGTCCACG	543
OY	587	PCysGluAspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgSerGlnHisAs	607
Db	544	CTGCAGAGATCTCAGACATGTACAAATACCCGCGTCTGCTCCGACAGGTGTGACATAGA	603
OY	607	PleuThr 609	
Db	604	CTTGACC 610	

RESULT 4.	
BB613947	
LOCUS	638 bp, mRNA, linear, EST 26-OCT-2001
DEFINITION	BB613947 BB613947 RIKEN full-length enriched, 0 day neonate head Mus musculus CDNA clone 4832415J17 5', mRNA sequence.
ACCESSION	BB613947
VERSION	BB613947.1 GI:16454417
KEYWORDS	EST.
SOURCE	house mouse.

ORGANISM	REFERENCE	AUTHORS
MUS MUSCULUS		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	1 (bases 1 to 638)	Arakawa, T., Carlini, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, T., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ono, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tadam, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y.
TITLE		
JOURNAL		
COMMENT		
		RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
		Unpublished (2001)
		Contact: yoshinhide Hayashizaki

Laboratory for genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by life technologies.
Note: this is a NIH-MGC library."

BASE COUNT 263 a 323 c 283 g 197 t
ORIGIN

Alignment Scores:

Pred. No.:	3-27e-84	Length:	1066
Score:	948.50	Matches:	197
Percent Similarity:	91.82%	Conservative:	5
Best Local Similarity:	89.35%	Mismatches:	16
Query Match:	24.82%	Indels:	4
DB:	12	Gaps:	0

US-09-836-392-21 (1-728) x BF796030 (1-1066)

QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheserGluPheArgGlnGlu 20
DB 173 ATGCTGAGGACCTGGCGGCGACCGATGCGATGAGAACCTTCTCCAGTTCGGCGAGAG 232
QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
DB 233 GCCAGCATCTCGACCGCGCTGCAGCAGCCCTGATGCTGCGCTCATCGCATCAGCATC 292
QY 41 HisPro-LeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSe 60
DB 293 CACCGGACTGCTGCTGCGCTGAGGCTGCGCGCTCAGCAGCGCTCAGCAGCGTGTGTC 352
QY 60 rGluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAl 80
DB 353 CGAGAGACCGCAAGATCTTCTTATACCCCTGCGACACATGCTCACCAAAATATAC 412
QY 80 ATyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLe 100
DB 413 CTRCCAGATCGCTCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 472
QY 100 uLysSerAspAsnIleLeuAlaTyrSerLeuAspValLysGlnHisIleAsnIleLysLe 120
DB 473 GAAGTCGAGACACATCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532
QY 120 uSerAspTyrGlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnGly 140
DB 533 ATCTGCTACGCGGATTTTCAGGCGACATTCATTCAGGCGCGCTGAGGCGGCGAC 592
QY 140 rProGlyTyrGlnAlaProGlnIleArgProArgIleValTyrAspGlnLysValAspHe 160
DB 593 TCTGCTACCGAGCGCCCGAGATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652
QY 160 rPheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGln 180
DB 653 TCTTCTCTTGAATGAGTGTCTCTAGAGTGTCTGCTGAGACAGGCGCTGCTGAGCCA 712
QY 180 sHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeu-GlyGlnP 200
DB 713 CCAGAGCTTCGATTCAGACAGAGGTGTC--CAAGGACATCCGCGCGTACTGAGACAGC 770
QY 200 rGlnGlnValGlnPheArgArgGlnAlaLeuMetGlnGlySTPAspThr 218
DB 771 CGAAGGAAGCTGCATTCGCGCATGCTGCTCA--TTATTGAGATCTGCGAAGCT 825

RESULT 6
LOCUS BC018934 1835 bp mRNA linear HTC 11-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:4120491, mRNA.
ACCESSION BC018934
VERSION BC018934.1 GI:17511956
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (ILN).
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigrl.nih.gov
Shevchenko, Y., Wetherby, R.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Leisapt, R.,
Lim, M., Maduro, O.L., Maeliello, C., Maestrian, S.D., McLooney, J.C.,
McDowell, J., Pearson, R., Snyder, B., Startip, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
Series: IRAL Plate: 20 Row: C Column: 22
This clone has the following problem: Incomplete processing.

FEATURES

source
1..1835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4120491"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_id="NIH-MGC-17"
/lab_host="DH10B-R"
/note="Vector: pCRB7"
BASE COUNT 446 a 493 c 475 g 421 t
ORIGIN

Alignment Scores:

Pred. No.:	2.94e-83	Length:	1835
Score:	943.00 <td>Matches:</td> <td>181</td>	Matches:	181
Percent Similarity:	99.45% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	99.45% <td>Mismatches:</td> <td>1</td>	Mismatches:	1
Query Match:	24.67% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	11	Gaps:	0

US-09-836-392-21 (1-728) x BC018934 (1-1835)

QY 547 SerLeuThrAspTyrCysSerMetSerTyrSerSerSerProProArgGlnAlaAla 566
DB 11 TCACACACGACTACTGCTCCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 70
QY 567 ArgSerProSerSerLeuProSerSerProAlaSerSerSerSerValProPheSerThr 586
DB 71 AGGTCCCTCCAGAGCTCCCGAGCTCCCGAGAGATTTCTTCAGAGTGTGCTTCTTCCACC 130
QY 587 AspCysGluAspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgSerGlnHis 606
DB 131 GACTCGAGGAGTACAGACATGCTACATACGCCGGGTGCTCGAGACAGGTGAGCAT 190
QY 607 AspLeuThrProMetAspGlyGluThrPheSerGlnHisLeuGlnAlaValLysIleLeu 626
DB 191 GACCTGACCCCATGAGAGGAGGAGACCTTCACCCAGCAGCTGAGCGGTGAAGATCTTC 250
QY 627 AlaValArgAspLeuIleThrValProArgArgGlyGlyAspValIleValIleGlyLeu 646
DB 251 GCCGTACAGACCTCATTTGGTCCCGAGCGCGGTGAGATGTTATGCTATTGGCTGG 310

source
1..631
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="X1032a19"
/clone_lib="NIBB Mochi normalized Xenopus neurula library"
/tissue_type="whole embryo"
/dev_stage="stage 15"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; CDNAS were oligo-dT primed and directionally cloned. Staging according to Neuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

BASE COUNT 156 a 157 c 150 g 167 t 1 others

ORIGIN

Alignment Scores:
Pred. No.: 4,42e-79 Length: 631
Score: 893.50 Matches: 172
Percent Similarity: 89.95% Conservative: 16
Best Local Similarity: 82.30% Mismatches: 20
Query Match: 23.38% Indels: 1
DB: 13 Gaps: 1

US-09-836-392-21 (1-728) x Bf042551 (1-631)

OY 28 GlnHisProCysIleValAlaLeuIleGlyIleSerIleHisProLeuGlySerPheAlaLeu 47
DB 3 CGGCACGAGGCGATGTGATCCCTGATGCGATCAGCATCACCACCTGCTGGCCCTA 62
OY 48 GluLeuAlaProLeuSerSerLeuAsnThrValLeuSerGluAsnAlaArg---AspSer 66
DB 63 GAGCTGGACCTCTTGGGGCCTCAACACAGTTCTGGGACACACCCCACTGCACT 122
OY 67 SerPheIleProLeuGlyHisMetLeuThrGlnIleAlaIleAlaSerGly 86
DB 123 TCATTATGCTCTGTGATGATGCTCACTACGCGGTTGCTATCAATAGCTTGTGA 182
OY 87 LeuAlaIleGlyLeuHisLysLysAsnIleIlePheCysAspLeuLysSerAspAsnIleLeu 106
DB 183 CTTTCTTACCTTCATATAAAGATATATATTTGGACCTGAAATCTGATTAACATTCG 242
OY 107 ValTrpSerLeuAspValLysGluHisIleAsnIleLysLeuSerAspTrpGlyIleSer 126
DB 243 GTTTGGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
OY 127 ArgGlnSerPheHisGluIleValAlaLeuGlyValGluGlyThrProGlyThrGlnAlaPro 146
DB 303 CGGCTGTCTATTCATGAGAGAGACTCTGGGTGGAGAGAGAGAGAGAGAGAGAGAGAGACT 362
OY 147 GluIleArgProArgIleValAlaIleGlyAspGluLysValAspMetPheSerTrpGlyMetVal 166
DB 363 GAAATCCCGCTCCGGGTGTGTATGATGATAAAGGTGATATGTTCTCATATGGAATGTC 422
OY 167 LeuTrpGluLeuLeuSerGlyLysIleArgProAlaLeuGlyHisIleGlnLeuIleAla 186
DB 423 CTTGATGAGCTATTTCTCAGGCCAGAGACCATTTGGGGCAGACCCAGGATGAGANTTCC 482
OY 187 LysLysLeuSerLysGlyIleArgProValLeuGlyLysIleGlnIleValAlaGlnPheArg 206
DB 483 AAGAGAGCTGTCAGAGGGATGTCAGATCTGGAGCGCCAGAGAGAGAGAGAGAGAGAGAG 542
OY 207 ArgLeuGlnAlaLeuMetGluCysTrpAspThrLysProGluLysArgProLeuAla 226
DB 543 CATCTACAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
OY 227 LeuSerValIleSerGlnMetLysAsp 235
DB 603 GCATCAGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 629

RESULT 9
AL702976 737 bp mRNA linear EST 22-MAR-2002
LOCUS AL702976

DEFINITION DKFZp686M2317_r1 686 (synonym: h1cc3) Homo sapiens cDNA clone
ACCESSION DKFZp686M2317.5, mRNA sequence.
VERSION AL702976.1 GI:19686331
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 737)
Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann
S.
EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Duesterhoeft A
MIPS
Am Klopferstr. 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686M2317) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..737
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="DKFZp686M2317"
/clone_lib="686 (synonym: h1cc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplex2; Site_1: SfilA; Site_2: SfilB;
CDNA-collection"

BASE COUNT 160 a 198 c 212 g 166 t 1 others

ORIGIN

Alignment Scores:
Pred. No.: 5.8e-78 Length: 737
Score: 883.50 Matches: 175
Percent Similarity: 71.43% Conservative: 0
Best Local Similarity: 71.43% Mismatches: 1
Query Match: 23.12% Indels: 69
DB: 9 Gaps: 1

US-09-836-392-21 (1-728) x AL702976 (1-737)

OY 219 LysProGluLysArgProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPhe 238
DB 3 AAGCAG 62
OY 239 AlaThrPheMetTrpGluLeuGlyCysGlyLysGlnThrAlaPhePheSerSerGlnGly 258
DB 63 GCCACCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
OY 259 GlnGluTrpThrValValPheTrpAspGlyLysGluGlnIleArgAspTrpThrValVal 278
DB 123 CAGGAGTACACCGCTGTGTTGGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182
OY 279 AsnThrGluLysGlyLeuMetGluValAlaGlnArgMetCysCysProGlyMetLysValSer 298
DB 183 AACACAG 242
OY 299 CysGlnLeuGlnValAlaGlnArgSerLeuTrpThrAlaThrGluAspGlnLysIleTrpIle 318
DB 243 TCCAGCTCCAGGTCAGAGATCCCTGTGGACAGCAGCAGGAGAGAGAGAGAGAGAGAGAG 302
OY 319 TyrThrLeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAla 338
DB 303 TACACCTCAGAGAGAGATGTGCCCCCTTAACACACACCCCAACAGGAGAGAGAGAGAGAG 362

LOCUS	BF583283	864 bp	mRNA	linear	EST 12-DEC-200
DEFINITION	602101746F1 NCI_CGAP_C024 Mus musculus cDNA clone IMAGE:4224922 5				
ACCESSION	BF583283				
VERSION	BF583283.1				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus				
AUTHORS	1 (bases 1 to 864)				
TITLE	NIH-MGC http://imgc.ncl.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection. (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-r@mail.nih.gov				
	Tissue Procurement: Jeffrey E. Green, M.D.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	http://image.llnl.gov				
	Plate: LLAM9815 row: 1 column: 11				
	High quality sequence stop: 630.				
FEATURES	Location/Qualifiers				
Source	1..864				
	/organism="Mus musculus"				
	/strain="FVB/N"				
	/db_xref="taxon:10090"				
	/clone="IMAGE:4224922"				
	/clone_1lb="NCI_CGAP_C024"				
	/lab_host="DH10B (T1 phage-resistant)"				
	/note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI;				
	Site_2: SalI; cloned unidirectionally. Primer: oligo dt				
	Average insert size 1.6 kb. Constructed by Life				
	Technologies. Note: this is a NCI_CGAP Library."				
BASE COUNT	207 a 247 c 238 g 172 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	2,71e-77	Length:	864		
Score:	878.00	Matches:	199		
Percent Similarity:	75.968	Conservative:	19		
Best Local Similarity:	69.348	Mismatches:	60		
Query Match:	22.978	Indels:	12		
DB:	12	Gaps:	5		
US-09-836-392-21 (1-728) x BF583283 (1-864)					
QY 345 AlAValProValIleIyLysIySAsnSerIyLeuValIleuAlagIyLeuAlaSpIyIeu 364					
2 GCACTACCTGTATCAAAAGAACTCCTTCCTGCTGCTGCAGGCGTGCATGCACTGC					
QY 365 ValAlaValaPheProValValaIrgIyTTProlYAspSerCySserIyIeuCySser 384					
62 GTGGCTGTGTTCCTGTGGCAGCGGGCACCCCAAGAAAGAGCTGCTCCTATCTGTCTCC					
QY 385 HisThrIaAsnArSerIySPheserIleAlaAspGluAspAlaIrgIAsnProTyr 404					
122 CACACACCCAAACAGTCCAAAGTCTGCAATCCCGAGAAAGATGACAGGGAGAACCTTAC					
QY 405 ProValIyAlaMetGluValValaAsnSerGlySerGluValATrPTySerAsnGlyPro 424					
182 CCACTGAAGCAATGAAAGTGTCACAGAGGGGTCTGAGGTCTGGTACAGCAAGGGCT 241					
QY 425 GlyIeuIeuValIleAspCySAlaSerIeuGluIleCySArgIrgIeuIuProTyrMet 444					
242 GGGCTCTCTGCTACGCTGCAACATCTCCACATATGACAGGCGGCTGAGGCTTACCA 301					
QY 445 AlaProSerMetValThrSerValValIySerSerGluGlyArgGlyGluIuValVal 464					

Db 302 GCCTGGTCCATGCTGACGCTCACTTGTGTGACAGCTCAGACTGCAGAGGAGGAGGAGTGGTC 361

QY 465 TTPCYSLAASPAPSLYALALASNSerLeuValMetYrYHSSerThrTyrGlnLeu 484

Db 362 TGGTCCCTCGACGACAAAGGCCCAATGTCTTAGTGAATGACCACTCAGCCATCCAGAGCTC 421

QY 485 CysAlaAaGTYrPhrCysGlyValProSerProLeuValAspMetPheProValArgPro 504

Db 422 TGTGCCCCGGTACTTGTGGGGAGCCCAATCTCTCAGGAGACGTTTCTGTGAGCCT 481

QY 505 LeuAspTrpGluProAlaAlaSerHisThrAlaAsnProGlyValProGlnGlyAsp 524

Db 482 TCAGTCCCTGAAACCCCA---GGCAGTCAACAAC-ACCTCAAGGGCCCTGGAAAG 537

QY 535 SerIleAlaAspValSerIleMet-TyrSerGlnGlnLeuGlnGlnIleLeuIleIle 544

Db 538 TGCATCCGCGGATGTGAGATCATGTTCAGCGAGAGAGCTGGCAGCAGATCCT----CA 593

QY 544 SGInGlnSerLeuThrAspTrpCysSerMetSerGlySerSerSerProProArgI 564

Db 594 CCAAAGAGATGCTCAGCATACTGCTCCATGTC-TCTTATTTCTGCTCCCAACCCAGCA 652

QY 584 nAlaAlaArgSerProSerSer-LeuProSerSerProAlaSerSerSerValProp 584

Db 653 AGATCTTAGTCCCTCTTCCAGTCAAGCCC--AGTCCCTCAACAAGTATTCAGTGTCT 709

QY 584 heSerThrAspCysGlnAspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgS 604

Db 710 TCTCTGGCAGATATGAAAGATCAAAAGGATGGAGAGGAGGAGGAGGAGTCA-----CGNG 760

QY 604 eRGlnHisAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeuGlnAlaValI 624

Db 761 AAGAATGATGATGACTGAGACCAAGATGGAGAGAA--GCAGACTGGGTGGAGTGTGA 817

QY 624 ysIleLeuAlaValArg 629

Db 818 AAGCCAGGCGCCAGCGG 834

RESULT 12

BB649591

LOCUS

DEFINITION BB649591 RIKEN full-length enriched, 16 days embryo head Mus

ACCESSION BB649591

VERSION BB649591.1 GI:16483846

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 715)

AUTHORS Arakawa,T., Carolin,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Himoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M., and Hayashizaki,Y.

TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

JOURNAL Unpublished (2001)

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shohri-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carlincl,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new


```

DR  Genew: HGNC:3655; FER.
DR  MIM: 176942; -.
DR  InterPro: IPR001060; Cdc15_Fes_CIP4.
DR  InterPro: IPR000719; Euk_Pkinase.
DR  InterPro: IPR000980; SH2.
DR  InterPro: IPR001245; Tyr_Pkinase.
DR  Pfam: PF00017; SH2; 1.
DR  Pfam: PF00069; Gkinase; 1.
DR  PRINTS: PR00109; TYRKINASE.
DR  ProDom: PD000001; Euk_Pkinase; 1.
DR  ProDom: PD000093; SH2; 1.
DR  SMART: SM00055; FCH; 1.
DR  SMART: SM00252; SH2; 1.
DR  SMART: SM00219; TYRK; 1.
DR  PROSITE: PS0133; FCH; 1.
DR  PROSITE: PS0107; PROTEIN_KINASE_ATP; 1.
DR  PROSITE: PS00109; PROTEIN_KINASE_DOM; 1.
DR  PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR  TRANSFERASE: Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
KW  Phosphorylation; Nuclear protein; SH2 domain; Polymorphism.
FT  DOMAIN 1 58 FCH.
FT  DOMAIN 460 550 SH2.
FT  DOMAIN 563 816 PROTEIN KINASE.
FT  NP_BIND 569 577 ATP (BY SIMILARITY).
FT  BINDING 591 591 ATP (BY SIMILARITY).
FT  ACT_SITE 684 684 BY SIMILARITY.
FT  MOD_RES 714 714 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT  VARIANT 439 439 V->L.
SO  SEQUENCE 822 AA; 94623 MW; 0491CD69392DB415 CRC64;

Query Match
Best Local Similarity 33.0%; Score 231.5; DB 1; Length 822;
Matches 73; Conservative 35; Mismatches 82; Indels 31; Gaps 10;

QY 16 EFROEASMLALQHPICVALIGISI--HPLCFALFLAPLSSIMTYLSENARDSSFIPLGH 73
DB 604 KFLQEAHKLKQYDHPNIVKLGICTQRPYIIMELVSGDFLFL--RRKDELKIKQ 660
QY 74 MLQKLIYQIASGLAVIAHKNITFCDLKSDNLIWMSLDVKEHINIKLSDYGISRQSPHEG 133
DB 661 LV--KESLDAAAGMLYDESKNCIHRDLARNCLVGENNV-----LKISDFGMSRQ--EDG 711
QY 134 ALGVEGTPG-----YQAEIRIRIYDEKVMFSYGMVLYELLS--GQRPALGHQLOI 185
DB 712 --GVYSSSGKQIPIKWTAPALNYGRYSSESDVWSFGILMETFSLGVCYPQMTNOQA 769
QY 186 AKRLSKGIRPVLGO--PEEYQFRRLQALMECWDTKPERRP 224
DB 770 REQVERGYRMSAPQHCPEDI-----SKIMAKCDYKRENRP 805

```

Search completed: April 8, 2003, 10:23:43
 Job time : 37 secs

```

Db 190 MetGluPheCysAlaGlyGlnLeuTyrGluValLeuArgAlaGlyArg----- 206
QY 199 TCCTTTATACCCCTGGGACATGCTCCACCCCAAAATAGCTTACCATGCTCGGGC 258
Db 207 -----ProValThrProSerLeuLeuValAspTrpSerMetGlyIleAlaGlyGly 223
QY 259 CTGGCCCTACCTGCGACAGAAAAATCATCTTGTGTGATCTGAGCTGAGACATTTCTG 318
Db 224 MetAsnTyrLeuHisLeuHisIleLeuHisArgAspLeuLysSerProAsnMetLeu 243
QY 319 GTGTGTGCTCCCTGAGCTGACAGACATCAATCAATCAATCTGTGAGTACGGGATTTCC 378
Db 244 IleThrTyrAspAspVal-----ValLysIleSerAspPheGlyThrSer 258
QY 379 AGGCAAG---TCATTCATGATGAGGGCCCTGAGGCGATGAGGACATCTGGTACACAGCC 435
Db 259 LysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThrValAlaTrpMetAla 278
QY 436 CCAGAGATCAGGCGCTCGCATTTGTATATGATGAGAAAGATGAGATATGTTCTCTATGGAGT 495
Db 279 ProGluValIleArgAsnGluProValSerGluLysValAspIleThrPheGlyVal 298
QY 496 GTGCTCTAGAGATTTGCTGACAGACAGCGCCCT----- 528
Db 299 ValIleThrGluLeuLeuThrGlyGluIleProTyrLysAspValAspSerAlaIle 318
QY 529 -----GCACCTGGGCGCACACAGCTCCAGATTGCCAAGAGCTGTCAAGGCAATCCGC 582
Db 319 IleTrpGlyValGlySerAsnSerLeuHisLeu----- 329
QY 583 CCGGTTCTGGGGGACCGCGGAGAGATGCACTTCGGCGACTGCGAGCGCTCATGATGAG 642
Db 330 ProValProSerSerCysProAsp-----GlyPheLysIleLeuLeuArgGln 345
QY 643 TGTGGGACACATGAGGACAGAGAGACGCG-----CTGGCCCTGTGG 684
Db 346 CysTrpAsnSerLysProArgAsnArgProSerPheArgGlnIleLeuLeuHisLeuAsp 365
QY 685 GTGTGAGCCAGATGAAGACCCGACTTTTGGCCACTGTATGATGAGCTGTGGG 744
Db 366 IleAlaSerAla-----AspValLeuSerThr 374
QY 745 AACGACAGACCTTTCTTCATCCAGGCGCAG-----GAGTAC 783
Db 375 ProGlnGluThrTyrPheLysSerGlnAlaGluTrpArgGluValLysLeuHisPhe 394
QY 784 ACCGTGCTTTGGATGAG----- 804
Db 395 GlnLysIleLysSerGlnGluThrCysLeuHisArgLeuGlnGluLeuValMetArg 414
QY 805 ---AAAGAGAGATCCAGGAGC----- 822
Db 415 ArgArgGluLeuLeuArgHisAlaLeuAspIleArgGlnHisTyrGluArgLysLeuGln 434
QY 823 -----TACAGGTGTGTAACACA-----GAG 843
Db 435 ArgAlaAsnAsnLeuTyrMetGluLeuAsnAlaLeuMetLeuGlnLeuGluLysGln 454
QY 844 AAGGCGCTCATGAGAGTGCAGAGATGTGC-----TGCCTGGAGTGAAGGTGAGC 894
Db 455 ArgGluLeuLeuArgArgGlnAlaLeuGlnArgArgCysProGlyLeuLeuLys-Pr 474
QY 895 TGGCAGCTCCAGTCCAGATACCTGTGTGACAGCCAGCAGGACGACGAAATCTACATC 954
Db 474 OHAspProSerArgGlyLeuLeuHisGlyAsnThrMetGluLysLeuIleLysLysAspAs 494
QY 955 TACACCTCAAGGAGCATGTGCTTAAACACACCCACAGGCGCTGTGATCTCAGCT 1014
Db 494 nValProGlnAsnLeuSer-----ProHisSerGlnArgProAspIleLeuLysAl 511
QY 1015 GTCTGTACCTGCTTTGTGGCGGCTGTTATTTAAAGAAATTCATCTGCTGTATGCG 1074
Db 511 agLuser-----LeuLeuProLysLeuAspAlaAl 521

```

```

QY 1075 GACCTGCCGATGGGCTTTGTGGCTGTGTTCCCTGGTGGGCGGACCCCAAGACAGC 1134
Db 521 AleuSerGlyValGlyLeu----- 527
QY 1135 TGCCTCTACCTGTGCTACACAGACGCAAGGTCAGATTGACATGCGGATGAAGAC 1194
Db 528 ---ProLysCysProLysAlaProProSer---ProLysArgSerArgArgGlyLysTh 545
QY 1195 GCACGCGAGAACCCCTACCCAGTGAAGCCATGAGGTGTGTCAACAGCGGCTGTAGGTC 1254
Db 545 ArgHisArg-----LysAlaSerAlaLys 553
QY 1255 TGTACACCAATGGGCGCGGCTCTTGTATCATGACTGTGCTCCCTGAGATCTGCAAG 1314
Db 553 sGlySerCysGlyAspLeuProGlyLeuArgThrAlaValProHisGluProGlyGly 573
QY 1315 CGGCTGAGACCT-----ACATGGGCC-----CCTCATG 1344
Db 573 yProGlySerProGlyGlyLeuGlyGlyGlyProSerAlaTrpGluAlaCysProProAl 593
QY 1345 GTTACGT-----CA 1353
Db 593 AleuArgGlyLeuHisAspLeuLeuLeuArgLysMetSerSerSerProAspLe 613
QY 1354 GTCTGTGCTGAGCTGTGAGGCGCAGAGGAGAGTCTGTGTGCTGCTGG-----ATGAC 1407
Db 613 uLeuSerAlaAlaLeuGlySerArgGlyArgGlyAlaThrGlyGlyAlaGlyAspProGly 633
QY 1408 AAGGCAACTCTTGTGTATGTATACCATCTCCACCACCTACCATCTGTGTGCCGGT---AC 1464
Db 633 ySerProProProAlaArgGlyAspThrProProSerGluGlySerProProGlySerTh 653
QY 1465 TTCTGGGGGGTCCCGAGCC---CCCTCAGGG 1492
Db 653 rSerProAspSerProGlyGluProLysGly 663

RESULT 4
US-09-428-711A-16
; Sequence 16, Application US/09428711A
; Patent No. 6358720
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Masaaki
; APPLICANT: Shirasawa, Takuji
; APPLICANT: Tokumitsu, Hiroshi
; APPLICANT: No. 6358720uch1, Teruhisa
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE
; FILE REFERENCE: 06501-045001
; CURRENT APPLICATION NUMBER: US/09/428,711A
; CURRENT FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: PCT/JP98/01246
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: JP 9/124798
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1050
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-428-711A-16

Alignment Scores:
Score: 7.08e-12 Length: 1050
Percent Similarity: 261.00 Matches: 196
Best Local Similarity: 35.578 Conservative: 95
Query Match: 23.968 Mismatches: 311
DB: 4 Indels: 217
Gaps: 38

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-428-711A-16 (1-1050)
QY 55 CAGAGGCGCAGATGCTGACAGCGGCTGACAGCCCTGATCTGCGGCTCATGCGCATC 114

```

DB 62 LysGluIleLysIleLeuLysGluLeuLysHisGluAsnIleValAlaLeuTyrAspPhe 81
OY 115 AGC-----ATCCACCCGCTGCTTGCCTCGAGCTCGG 150
DB 82 GlnGluMetAlaAsnSerValTyrLeuValMetGluTyrCysAsnGlyLysPheAla 101
OY 151 CCGCTGAGAGCGCTCAACACCGCTGCTCGGAGAACGCCAGATTCTCTTATACC 210
DB 102 AspTyrLeuHisAlaMetAlaThrLeuSerGlnAspThr----- 114
OY 211 CTGGGACACATGCTCACCACCAAAATAGCTTACAGATCGCTCGGCGCTGACTCTG 270
DB 115 -----IleArgLeuPheLeuGlnGlnIleIleLeuAlaMetArgLeuLeu 129
OY 271 CACAGAAAATCATCTTCTGTGTGACTGAGTGGAGACATCTGTGTGTGCTT 330
DB 130 HisSerLysGlyIleIleHisArgAspLeuLysProGlnAsnIleLeuSerAsnPro 149
OY 331 GACGTCAAG-----GAGCACATCAACATCAAGTATCTGACTACGGGATTTGC 378
DB 150 AlaGlyAlaArgAlaAsnProAsnSerIleArgValIleAlaAspPheGlyPheAla 169
OY 379 AGG-----CACTCATTCATGAGGCGCCCTGAGCGCTGAGGACCTCGCTACCA 432
DB 170 ArgTyrLeuGlnSer---AsnMetAlaAlaThrLeuCysGlySerProMetTyrMet 188
OY 433 GCCCCAGAGATCAGCGCTCGCATGTATATGATGAGAGATGATATGTTCTCTATGA 492
DB 189 AlaProGlnValIleMetSerGlnHisTyrAspGlyLysAlaAspLeuTyrSerIleGly 208
OY 493 AAGGCTCTAGAGATTGTGTGCAGACAGCGCCCT-----GACTGGGCGACCAACAG 546
DB 209 ThrIleValTyrGlnLysLeuThrGlyLysAlaProPheGlnAlaSerProGlnAsp 228
OY 547 CTCACAGATTGCCAAGAGCTGTGTCCAGACATCCGCCGCTTGGGGCGCGGAGAA 606
DB 229 LeuArgLeuPheTyrGluLysAsnLysThrLeuValProThrIleProArgLysThrSer 248
OY 607 GTGCACTTCGGCGACTGCAGCGCGCTCATGATGAG----- 642
DB 249 AlaProLeuArgGlnLeuLeuAlaLeuGlnArgAsnHisLysAspArgMetAsp 268
OY 643 -----TCCTGGACACTAACCCAGACAGAGCGCGCTGGCC 678
DB 269 PheAspGlnPhePheHisIleProPheLeuAspAlaSerProSerValArg----- 285
OY 679 CTGTCCGTTGTGAGCCAGATGAAGACCGCATTTGGCA---CCTTCATGATGAACAG 735
DB 286 -----LysSerProPro-ValProValProSerTyrProSerSer 298
OY 736 TGTGTGGGAGACAGACCGCTTCTTCATCCAGGCGGAGAGTACACCGTGTGTTT 795
DB 298 rGlySerGlySerSerSerSerSerSerSerThrSerHisLeuAlaSerProSerSer 318
OY 796 TGGG-----ATGGAAAAG-----AGGACTCCAGAGAACATACAGGTTGTT---G 834
DB 318 uGlyGlnMetGlnGlnLeuGlnLysThrLeuAlaSerProAlaAspThr-AlaGlyPheAl 338
OY 835 AACACAGAGAGCGCTCATGAGTGCAGAGAT----- 869
DB 338 eHisSerSerArgAspSerGlyGlySerLysAspSerSerCysAspThrAspAspPhe 358
OY 870 ---GTGTGCGCTGGGAT---GAAGGTGAGCTG-----CCAGCTCCAGTCCAGAGATCC 918
DB 358 AlaMetValProAlaGlnPheProGlyAspLeuValAlaGlnAlaProSerAlaLysProp 378
OY 919 CTGTGGACAGC-----CACCCAGAGACCGAAATC 948
DB 378 rGlyProAspSerLeuMetCysSerGlySerSerLeuValAlaSerAlaGlyLeuGlnSer 398
OY 949 T-----ACATCTACACCTCAAGGCGCATGTGCCCTTAACACACCCCAACAGGCTT 1001

DB 398 IsGlyAlaThrProSerProSerProCysSer----- 409
OY 1002 GGATACTCCAGTGTGTCACCTGCTTCTTGGCCGCTGTTATTAAAGAAATTCCTA 1061
DB 410 -----SerSerProSerPro----- 414
OY 1062 CCTGTCTTACGGGCGCTCGCGATGGCTTGTGCTGTGTTCCGTTGCGGGGCGAC 1121
DB 415 -----SerGlyAlaGlnGlyProPheSer-----SerSerArgCysGlyAla- 428
OY 1122 CCCAAGACAGCTGCTCTTACCTGTGTCTACACACAGCCACAGGATCCAAAGTTCAGAT 1181
DB 429 -----SerValProIle-----ProValProThrGlnValGlnAsnTyrG 442
OY 1182 CCGGATGAGAGACGCGACGCGACAGACCCCTACCCAGTGAAGGCCATGAGGTGTCACAG 1241
DB 442 InArgIleGlnArgAsnLeuGlnSerProThrGlnPheGlnThrProArg---SerSerA 461
OY 1242 CCGCTCTGAGCTGTGTACAGA-----ATGGCGGGCGCTTCCTGTATGACTGTGC 1295
DB 461 lAlaArgArgSerGlySerThrSerProLeuGlnPheAlaArgAlaSerProSerProp 481
OY 1296 CTCCCTGAGATCTGCAGCGCGCTGAGCCCTACATAGCCCTCCATGTTACGTCAGT 1355
DB 481 rAlaHisAlaGlnHisGlyGly----- 488
OY 1356 CGTGTGACAGCTGTGAGCGACAGAGGAGAGAGTGTGTGCTGCTGATGACAGGCCAA 1415
DB 489 -----ValLeuAlaArgLysMetSerLeuGlyGlyAlaProTyrThrPro 505
OY 1416 CTCTCTGTGATGTACCACTCCACCACTACCAAGCTGTGTGCTGCTGCTTCTGCGGGT 1475
DB 505 eProGlnValGlyThrIleProGlnArg-----ProGlyTyrSer---GlyT 520
OY 1476 CCCCAGCCCTTCAGAGACATGTTCCCTGCGGCCCTTGGACAGGAGACCCCGGACAC 1535
DB 520 hrProSerProGlnGlnAlaGluMetArgGlyGly-----ArgSerProArgProG 537
OY 1536 CAGCAGACGCGCCAAACCAAGGTGCTGAGGGGAGCTCATCGGAGCGATGACATCAT 1595
DB 537 LysSerSerAlaProGlnHis-----SerProArgThrSerGlyLeuGlyC 552
OY 1596 GTACAGTAGAGAGCTGGGACCGACAGATCTGTATCCACAGGAATCATCTACTGACTACG 1655
DB 552 YArgLeuHisSerAlaProAsnLeuSer-----AspLeuHisValAla 567
OY 1656 CTCACATGCTCTCTACTCTCATCCACACCC---GCCAGGCTGGACAGGTCGCCCTGAG 1712
DB 567 rGProLysLeuProLysProProThrAspProLeuGlnValAlaPheSerProProGlnA 587
OY 1713 CCTCCCA----- 1720
DB 587 lAsrProProGlnProSerHisGlyLeuGlnSerCysArgAsnLeuArgLysSerProL 607
OY 1721 -----GCTCCCGACAGAG 1733
DB 607 YLeuProAspPheLeuGlnArgAsnProLeuProProIleLeuGlnLysSerProThrLysA 627
OY 1734 TTTCTCCAGTGTGCTTCTTCACAGCGATGCGAGAGCTCAGACATGCTACATACGCCGG 1793
DB 627 lValProSerPheAspPheProLysThrProSerSerGlnAsn-----LeuLeuAlaL 645
OY 1794 TGTGCTCCGACAGGTGTGACATGACCTGACCC----- 1828
DB 645 euleuAlaArgGlnGlnValMetThrProProArgAsnArgThrLeuProAspLeuS 665
OY 1829 -----CCATGAGAGGGAGACCTTACAGCCAGCACTGACAGCGCGGTAAAG- 1873
DB 665 eArgLysAlaGlyProPheHisGlyGlnProLeuGlyProGlyLeuArgProGlyGlnAspP 685
OY 1874 --TCTTCGCGCTCAGAGACCTATTTGGGTCCCAAGGCGCGGTGAGATGATGTGTCAT 1931
DB 685 rGlyGlyProPheGlnArgSerPheSerThrSerArgLeuThrAspLeuLeuLysA 705

```

OY 1932 TGGCCTG---AGAAGATTCTGAAGCCAGCGGCGGCGATTCGCTTAAAGC 1988
DB 705 laalaphelythrlnalaproasprioglyserthcluserleuglnlulysprom 725
OY 1989 CCGAGAGTACTCCGCGATGGGCTGCTGGATGCTG---CCGTCGTGCAAGAGCAC 2045
DB 725 eglulalelproseralaglypneglylserleuhsisproglalalaglyalg 745
OY 2046 TGTGTGTCACCTTGAAGAAATGAACACAGAGTGTGTCGCGCGCTGTGAGGGGCTG 2105
DB 745 lYthserseerProserProvalalpherthvalglyserProProserglyserthp 765
OY 2106 GGGCGCCAGGAGTTCGACA-----TTTTCTACAGCTCCTACGAGAGC 2149
DB 765 roProglndlyProargthnArgmetPheSerAlaglyProthrglyser 781

RESULT 5
US-09-428-711A-14
Sequence 14, Application US/09428711A
Patent No. 6358720
GENERAL INFORMATION:
APPLICANT: Muramatsu, Masaaki
APPLICANT: Shirasawa, Takuji
APPLICANT: Tokumitsu, Hiroshi
APPLICANT: No. 6358720uchl, Teruhisa
TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE
FILE REFERENCE: 06501-045001
CURRENT APPLICATION NUMBER: US/09/428,711A
CURRENT FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: PCT/JP98/01246
PRIOR FILING DATE: 1998-03-23
PRIOR APPLICATION NUMBER: JP 9/124798
PRIOR FILING DATE: 1997-04-28
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 1051
TYPE: PRT
ORGANISM: Mus musculus
US-09-428-711A-14

Alignment Scores:
Pred. No.: 1,18e-10 Length: 1051
Score: 245.00 Matches: 178
Percent Similarity: 32.92% Conservative: 87
Best Local Similarity: 22.11% Mismatches: 259
Query Match: 6.04% Indels: 281
DB: Gaps: 35

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-428-711A-14 (1-1051)
OY 55 CAGAGAGCCGACATGCTGCAGCGCTGCAGACCCCTGCATCGCGGCTCAGCGCATC 114
DB 62 lylselutleuylsleuLysgluleuLysHisglulubsnllevalalaleuLyrasphe 81
OY 115 AGC-----ATCCACCGCTGCTGCTCGCCCTGAGAGCTGCGC 150
DB 82 GlndluMetalaasenservatlYrleuValMetclulYrCysasnlglyAspleuAla 101
OY 151 CCGCTCAGACGACCTCAACACCGCTGCTCGAGAGCCGACAGATTCCTCTTATACC 210
DB 102 AspyrleuhsisThrMetArgThrleuSerclunspthr----- 114
OY 211 CTGGGACACATGCTCACCACAAATAAGCCACAGATCGCTCGGGCTGCGCTACTG 270
DB 115 -----ValArgleuPheleuGlnlnlealaglyAlaMetArgleu 129
OY 271 CACAAGAAAACATCATCTTCTGACCTGAAGTCGAGACATCTTGCTGCTGCTT 330
DB 130 HisserlysglyllelleHisArgAspleuLysProGlnsnlleleuLysAsnPro 149
OY 331 GACGTCAAG-----GAGCACATCAACATCAAGCTATGTGACTAGGATTTG 378

```

```

DB 150 GlylYAlaAlaAsnProserlnleArgvalLyslleAlaAspPheleuPhele 169
OY 379 AGC-----CACTCATTCATGAGGCGCCCTAGGCTGAGGAGCATCTCGCTACAG 432
DB 170 ArgYrleuGlnser---AsnMetleuAlaAlaThrleuCysGlySerProMetYrMet 188
OY 433 GCCCAGAGATCAGGCGCTGCAATTTGATATGAGAAAGTGAATATGTTCTCCATGA 492
DB 189 AlaProGlnVallleMetserGlnHisYrAspGlyLysAlaAspleuYrPserlle 208
OY 493 ATGTCCTTAGAGTTGTGTGCAGACAGCGCCCT-----GCACTGGGCCACACAG 546
DB 209 ThrIleValYrclnLysleuThrGlyLysAlaProPheGlnHisSerProGlnAsp 228
OY 547 CTCAGATTGCAAGAGCTGTCCAAGGAGCATCCGCGCTTCTGGGAGCGGAGGAA 606
DB 229 LeuArgleuPheYrclululysasnLysThrleuValProAlaIleProArglulThrser 248
OY 607 GTGCAGTTCGCGGACTGCAGCGGCTCATGATGAGAGTGTGGACACTAAGCCAGAG 666
DB 249 AlaProLeuArglndleuLeuAlaLeuGln-----ArgAsnHisLys 264
OY 667 -----CGACGCTGGCCCTGTGCGGTGAGCCAG 696
DB 265 AsparYMetAspPheAspLuphePheHisAspPheLeuAspAlaSerThrProIle 284
OY 697 ATGAAGAGCCGACCTTTGCA-----CCTTCATGATGAACCTGTGCTGGAGAGCAGA 753
DB 285 lylLysSerProPro-ValProValProserYrProserSerGlySerclYserSerse 304
OY 754 GCCTTCTTCATCCAGAGGCCAGAGTACACCGTGTGTTGGATGAAAAGAGAG 813
DB 304 rSerSerSer----- 307
OY 814 TCCAGGAACCTACACGCTGTGGAACACAGAGAAGGCGCTCATGAGAGTCAGAGATGTC 873
DB 307 ----- 307
OY 874 TGCCTGGATGAGGTGAGCTGCAGCTCCAGCTCCAGATCC-----TGTGAGAGCC 930
DB 308 -----SerAlaSerHisleuAlaSerProProserleuGlylume 321
OY 931 ACCGAGAGCCAGAAATCTACATCTACACCTTACAGGCGCATGCCCCCTTAAACACACC 990
DB 321 tProGlnleuGlnlylThrleuThrSerProAlaAspAlaAlagly----- 336
OY 991 CAACAGGCTTGATCTCAGCGTGTGCA----- 1021
DB 337 -----PheleuGlnLysSerArgAspSerGlylYserSerLysAspSerY 352
OY 1022 -----CCTGCTTGTGGCGCTGCTGCTTATTAAAGAT 1056
DB 352 sasprThrAspPheValMetValProAlaGlnPheProGlyAspLeuValAlaGlnAl 372
OY 1057 TCTTACCTGCTTACGGGCGCTCGCGATGGCTTGTGCTGTGTTCCCGTGTGCGG 1116
DB 372 aala-----SerAlaLysProProAspSerleu-----LeuCyse 385
OY 1117 GGCACCCCAAGAGACCT----- 1137
DB 385 rGlySerSerleuValAlaSerAlaGlyLeuGlnLysHisGlyYrArgThrProserProse 405
OY 1138 TCTTACCTGCTGCACACAG----- 1159
DB 405 rProthCysSerSerSerProserProserGlyArgProGlyProPheSerSerAsnAr 425
OY 1160 -----CACAAGCTCAAGTTCCACATCCGCGATGAA 1191
DB 425 gYrGlyAlaSerValProIleProValProthGlnValHisAsnYrGlnArgIleG 445
OY 1192 GAGCAGGAGAGAACCCCTCAGCGTGAAGGCCATGAGGTCGACAGGGCTCGAG 1251
DB ----- 1251

```

D	445	uglasmnleunglinsertProthrnglinslnthralatrg	-----SerserAlallleangAr	464
Q	1252	GTCTGGTACA	-----GCATGGGCGGCGCTCTTGTCTATGCAGTGTGCTCC	1299
D	464	gserdlserthtrProleugllypheyAlaagAlaserProser	-----ProPr	481
Q	1300	CTGGAGATCTGCAGCGCGGTGGAGCCCTACATGGCCCCCTCATGGTATACGTACAGTCTGG	1359	
D	481	oserHstthasply	-----	486
Q	1350	TGCAGCTCTTAGGCGACAGAGGAGAGAGTCTGTGCTGCTGGATGCAGAGCCACATCC	1419	
D	487	---AlamerleuAlarlgylsleuserleugllygly	-----ArgPro-----	500
Q	1420	TTGGTGAATGTCACATCCACCTACACAGACTGTGAGCCCGGACTTTCGGGGGCTCC	1479	
D	501	-----tytrProserProglInvaltlythrlleProglInlAargProserTtrPserAr	518	
Q	1480	AGCCCCCTCAGGACATGTTCCGTGCGGCCCTTGAGACAGAGAACCCCGGACCCAGC	1539	
D	518	gvalProser-----ProglnglyAlaasvalAargvalAglAargserPro-	533	
Q	1540	CACAGCGCCAAACCAAGGTCCTGAGGGAGATCCATGCGCGAGCTAGCATCATATAC	1599	
D	534	---ArProglYserSerValProgluHtserProAargthtrhlyleuglYcys-	551	
Q	1600	AGTAGAGAGCTGGGACAGCATCTGATTCACACAGATACACTACATGACTGCTCC	1659	
D	552	-----ArgleuHtser-----AlaPr	557	
Q	1660	ATGTCTCTTACTCTATATCCACACCCCGCAGGCTGCCAGGTCCCCCTCAAGCTCCC	1719	
D	557	oasnleuserasrphelshvalAargProYlsleuProYlsPro-----	572	
Q	1720	AGCTCCGCCAGCAGATTTCTCAGTGTGCGCTTCTCCAGC---ACTGAGAGACTCAGAC	1776	
D	573	---ProthrasrProleugllyAlathrPheaserProProglInthSerAlaProglInPr	591	
Q	1777	ATGCTACATACGCCCGGTGCTGCTCCCGACAGGTCTGACACATGACCTGACCCCATGAC	1836	
D	591	ocys-----	592	
Q	1837	GGGAGACCTTCAGCCAGCACCTGCAGAGCCGTGAGATCTCGCG	1882	
D	593	-----ProglYleuglInserYsarYgProleuAlrglyserProYlsleuProAsph	610	
Q	1883	-----TCAGAGACCTCATTTGG-	1900	
D	610	eIeuglInArYserProleuProProIleleuglYserProthrlYsAlaglyProserPh	630	
Q	1901	-----TCCCAGGCGCGGTGGAGAGATGTATATGTCATTTGGCTGTGAGAGG-	1945	
D	630	eAsrPherProlythtrProserSerGlInasleuLeuThrlleuLeuAlaYrglInglyA	650	
Q	1946	-ATTCTGAAGCCCAAGCGGGCGAGTCATTCGCCGTTTAAAGCCCGAGAGACTGACTCG	2004	
D	650	lvalMerthtrProbrAargAsnAtrghrlleuProAspleuserglAlaser	667	
Q	2005	CATGGGGTGTGGATGCTGCGCGTGGTGGCAAGAGACACTGTGTGTGTGCACCTTTGAA	2064	
D	668	-----ProPhenlsaglyInglInleu-----	674	
Q	2065	AATGAACAACAGAGTGTGCTGCGCGTGTGAGGGGCTGGGGGCGCAGAGATTGAC	2124	
D	675	-----glySerlyleuAargProAlaaglInasprhAargglyProPheglAargse	691	
Q	2125	ATTTTCTACAGT 2137		
D	691	rPheSerThrSer 695		

```

      : Patent No. 5843749
      : GENERAL INFORMATION:
      : APPLICANT: Maisonnier, et al.
      : TITLE OF INVENTION: ENK AND FOR TYROSINE
      : TITLE OF INVENTION: KINASES
      : NUMBER OF SEQUENCES: 107
      : CORRESPONDENCE ADDRESSES:
      : ADDRESSEE: Regeneron Pharmaceuticals, Inc.
      : STREET: 777 Old Saw Mill River Road
      : CITY: Tarrytown
      : STATE: NY
      : COUNTRY: U.S.A.
      : ZIP: 10591
      : COMPUTER READABLE FORM:
      : MEDIUM TYPE: Diskette
      : COMPUTER: IBM Compatible
      : OPERATING SYSTEM: DOS
      : SOFTWARE: FastSeq Version 2.0
      : CURRENT APPLICATION DATA:
      : APPLICATION NUMBER: US/08/469,537A
      : FILING DATE: 06-JUN-1995
      : CLASSIFICATION: 435
      : PRIOR APPLICATION DATA:
      : APPLICATION NUMBER: USSN 08/406,247
      : FILING DATE: 17-MAR-1995
      : APPLICATION NUMBER: USSN 08/144,992
      : FILING DATE: 28-OCT-1993
      : APPLICATION NUMBER: USSN 07/336,559
      : FILING DATE: 26-JUL-1991
      : ATTORNEY/AGENT INFORMATION:
      : NAME: Kempler, Ph.D., Gall M
      : REGISTRATION NUMBER: 32,143
      : REFERENCE/DOCKET NUMBER: REG 070C
      : TELECOMMUNICATION INFORMATION:
      : TELEPHONE: 914-345-7400
      : TELEFAX: 914-345-7721
      : TELEX:
      : INFORMATION FOR SEQ ID NO: 72:
      : SEQUENCE CHARACTERISTICS:
      : LENGTH: 800 amino acids
      : TYPE: amino acid
      : STRANDEDNESS: single
      : TOPOLOGY: unknown
      : MOLECULE TYPE: peptide
      : US-08-469-537A-72

Alignment Scores:
Pred. NO.:          1.26e-10           Length:          800
Score:              244.00             Matches:         100
Percent Similarity: 37.56%             Conservative:    66
Best Local Similarity: 22.62%            Mismatches:     157
Query Match:        6.01%               Indels:         119
DB:                  2                   Gaps:          15

US-09-836-392-8_COPY_22_2205 (1-2184) x US-08-469-537A-72 (1-800)

OY   4 CTGAGGACACTGGCGGGCCACGATGCATGAAGAATTCTCCGAGTCCSGGAGAGAGGCC 63
Db   11clyshtnleuyluaspytlaasnaaprobinsintrmetgtsuprheinglnlualaa 387

OY   64 AGCATGTCGACAGCGGCTCCACACGACCCCTGATCGTGCGGTGACATGACGCG-----ATCAGC 117
Db   388 SerleuMetalaigluLeuInhiNhiStrpAsnIleValcysLeuLeugAlaValThgin 407

OY   118 ATCCACCGCGCTGC-----TTGCGC 138
Db   408 GluInprValcyuMetleuRheglutytIlleasnclngluArleuinisglurPheleu 427

OY   139 CTGAGAGTCGGGCGGCTGACAGGCTCAACAACGCTGCTGSCGAGAACGACAGATTC 198
Db   428 IleMetaGserTronHisSerArvalaGlucysSerSerArspGluAsrgLytHValuys 447

OY   199 TGCTTAATACCSTGGGACACATGCTCACCCCAAATAAGCTATACAGATGAGCTCGGCG 258

```

```

Db 448 SerSerLeuAspHisGlyAspPheLeu-----HisIleAlaIleGlnIleAlaIleGly 465
QY 259 CTGGCCATCTGCACAGAAAAACATCATCTTCTGTGACCTGAGTCCGACACATCTTG 318
Db 466 MetGluTyrLeuSerSerHisPhePheValHisLysAspIleAlaIleAlaArgSniLeu 485
QY 319 GTGTGTCCTTGCAGTCAAGACACATCAACATCAAGCTATCTGACTAGGGATTGCG 378
Db 486 -----IleGlyGlnIleuHisValIleSerHisPheLeuLeuSer 500
QY 379 AGCGATCATTCATGAGGCGCCCTAGCGGTGAGGCACTCT-----GGC 426
Db 501 ArgGluIleTyrSerHisAspTyrTyrArgValGlnSerLysSerLeuLeuProIleArg 520
QY 427 TACGAGCCCGCAGATCAGGCGCTGATATGATGAGAGATGATGTTCTCC 486
Db 521 TrpMetProGluAlaIleMetTyrGlyLysPheSerSerAspSerAspIleTrpSer 540
QY 487 TATGGAATGCTCTAGAGTGTCTGCA---GGACAGCGCCCTGCACTGGGCGCACAC 543
Db 541 PheGlyValValLeuTrpGluIlePheSerPheGlyLeuGlnProTyrTyrIlePheSer 560
QY 544 CAGCTCCAGATTGCCAAGAGCTGTCCAGAGGCAATCCGCGTCTGTGGGCGAGCGGAG 603
Db 561 AsnGlnIleValIleGluMetValArgLys-----ArgGlnLeuLeuProCysSerGln 578
QY 604 GAGTGCACCTTCCGCGACTGCAGCGCCCTCATGATGATGATGCTGGGACACTAAGCCAG 663
Db 579 AspCysPro---ProArgMetTyrSerLeuMetTrpGluCysTrpAsnGluIleProSer 597
QY 664 AAGCGACGCGTGGCCCTGTGCGGTGAGCGAGATGAGACCGCCGCTTTCGACACTTC 723
Db 598 ArgArgProArgPheLysAspIleHisValArgLeuArgSer----- 611
QY 724 ATGTAAGACTGTGCTGTGGAGAGACAGACGCTTCTCTCATCCAGGCGCAGAGTAC 783
Db 611 ----- 611
QY 784 ACCGCTGCTTTTGGATGCA-----AAAAGAGAGTCCAG 819
Db 612 -----TrpGluIleuSerHisThrSerSerThrThrProSerGlyGly 627
QY 820 AACCTACGCGTGTGAACAGAGAGGCGCTCATGAGTGTGAGAGATGCTGCTCCCT 879
Db 628 AsnAlaThrThrGlnThrThrSerLeuSerAlaSerProValSerAsnLeuSerAsnPro 647
QY 880 GCGATGAAGTGAAGTCCAGCTCCAGTCCAGATCCCTGTGAGACACACCGAGAGAC 939
Db 648 -----Arg 648
QY 940 CAGAAATCTACATCTACACCTCAAGGCGATGCGCCCTTAACACACCCCAACAGGCC 999
Db 649 TyrProAsnTyrMetPheProSerGlnGlyLe-----ThrProGln----- 662
QY 1000 TTGGATACATCCAGCTGTGCACCTCTGTGCGCGTGTCTTAAAGAAATTC 1059
Db 663 -----GlyGlnIleAlaGlyPheIleGlyProProIleProGlnAsnGlnArg 678
QY 1060 TACCTGCTTACGCGGCGCTCGCGATGCGGCTGTGCTGTTCCTCCGTGTGCGGGCG 1119
Db 679 PheIleProIleAsnGlyTyr-ProIle----- 687
QY 1120 ACCCCAAAGAGAGCTGTCTTACTGTGTCTACACACACAGCCAAACAGT----- 1168
Db 688 -ProGluGlyTyrAlaAlaPheProAlaAlaHisTyrGlnProThrGlyProProArgVal 707
QY 1169 -----CCAAGTTCAGATGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1209
Db 707 LileGlnHisCysProProLysSerArgSerProSerSerAlaSerGlySerThrSe 727
QY 1210 TACC 1213

```

```

Db 727 rThr 728
RESULT 7
US-08-469-537A-78
Sequence 78, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisonnier, et al.
TITLE OF INVENTION: EHK AND FOR TYROSINE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempier, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 800 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-537A-78
Alignment Scores:
Pred. No.: 1,26e-10 length: 800
Score: 244.00 Matches: 100
Percent Similarity: 37.56% Conservative: 66
Best Local Similarity: 22.62% Mismatches: 157
Query Match: 6.01% Indels: 119
Gaps: 15
Db:
US-09-836-392-8_COPY_22_2205 (1-2184) x US-08-469-537A-78 (1-800)
QY 4 CTGAGCACCCTGGGCGCCAGATGCGATGCAACTTCTCCGAGTTCGCGGACAGAGGCC 63
Db 368 IleLysThrLeuLysAspTyrAsnAsnProGlnGlnTrpMetGluPheGlnGlnAla 387
QY 64 AGCATGTCGACGCGCTGCAGACACCCCTGCATGTGCGCTCATGGC-----ATCAGC 117
Db 388 SerLeuMetAlaGluLeuHisHisProAsnIleValCysLeuLeuGlyAlaValThrGln 407
QY 118 ATCCACCCGCTCTGC----- 118
Db 408 GluGlnProValLysMetLeuPheGluTyrIleAsnGlnCysPheuHisGluPheLeu 427
QY 139 CTGAGCTCGCGCGCTGCAGACAGCTCAACACCGTGTCTGCCGAGAGAGATTC 198

```

Db 406 TCCGATAGGAGGAGTGAATGACTTCGGCATCCCAAGAGCTGGGCATGAGGCGCTC 465
QY 388 TTCATAGAGGCGCCCTAGGCGGTGAGGCACTCTGCGTACCAGGCCAGAGATCAG 447
Db 466 ACCGAGCAGGCA-----TGGCGGTGGAAAGCGCGTGTACATGTCCCGGAGCAGGTG 519
QY 448 CTGCGATTTGTAATGATGAGAGGTAAGATATGTTCTCTATGGAATGTGCTCTACGAG 507
Db 520 ACGGCGGTGCGGTGATGCGCGCACCGACATCTTCTCGGCGTGTGCTCTACGAA 579
QY 508 TTGCTGACAGACAGCGCCCTGCACTGGGCCACCAAGCTCCAGATTGCCAAGAAGCTG 567
Db 580 GCCCTCTCCGCGCGCGCCCTTCCACGGCAAGAGCGGCGAGGTCTTCGGAAGATT 639
QY 568 TCCAAGGCGATCCGCGGTCTGCGGCGAGCGGAGAAATGCAAGTTCGCGCACTGCAG 627
Db 640 CGCGACGCGAAGTACACGCGCTCTCCAGGTGGCGGCCCAACGTCCCGCGCGCTGTG 699
QY 628 GCCTCATGATGAGTGTGAGCACTAAGCCAGAGAAAGCAGCGCTGGCCCTGTGCGGTG 687
Db 700 CGCATCATCCAGCGCGCATGAGGTGAAGCGGAGGACCGCTTCCCGAGCGCGCGCC 759
QY 688 GTGAGCC 694
Db 760 ATGCGCC 766

Search completed: April 15, 2003, 06:59:12
Job time : 1763 secs

[illegible]

```

Db      707  11lelnhtscyspPropoProlySserArgSerProserSerAlaSerGlySerThruse 727
QY      1210  TACC. 1213
Db      727  rThr 728

RESULT 8
US-08-469-537A-105
: Sequence 105, Application US/08469537A
: Patent No. 5843749
: GENERAL INFORMATION:
: APPLICANT: Maisondiplerie, et al.
: TITLE OF INVENTION: ENK AND FOR TYROSINE
: TITLE OF INVENTION: KINASAS
: NUMBER OF SEQUENCES: 107
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Regeneron Pharmaceuticals, Inc.
: STREET: 777 Old Saw Mill River Road
: CITY: Tarrytown
: STATE: NY
: COUNTRY: U.S.A.
: ZIP: 10591
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/469, 537A
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: USSN 08/406,247
: FILING DATE: 17-MAR-1995
: APPLICATION NUMBER: USSN 08/144,992
: FILING DATE: 28-OCT-1993
: APPLICATION NUMBER: USSN 07/736,559
: FILING DATE: 26-JUL-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Kempler, Ph.D., Gail M
: REGISTRATION NUMBER: 32,143
: REFERENCE/DOCKET NUMBER: REG 070C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 914-345-7400
: TELEFAX: 914-345-7721
: TELEX:
: INFORMATION FOR SEQ ID NO: 105:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 937 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FRAGMENT TYPE: internal
: FEATURE:
: NAME/KEY: Human ROR1
: LOCATION: 1...937
: OTHER INFORMATION:
US-08-469-537A-105

Alignment Scores:
Pred. NO.: 1.34e-10 Length: 937
Score: 244.00 Matches: 100
Percent Similarity: 37.56% Conservative: 66
Best Local Similarity: 22.62% Mismatches: 157
Query Match: 6.01% Indels: 119
Db: 2 Gaps: 15

US-09-836-392-8_COPY_22_2205 (1-2184) x US-08-469-537A-105 (1-937)
QY      4  CTGAGGCACCCGCGGGCCACCGATGCATGAAGAATTCTCCGAGTCCGCGGACGAGCC 63
Db      505  11elystrhrbulysAsPtyrAsnAsnProGlnIntrPwctGluPngInGlnGluAla 524

```



```

QY 64 AGCATCTGACGCGCTGACGACACCCCTGATCGTGGCGCTCATCGGC-----ATCAGC 117
   |||:||||| ||| ||||| ||||| |||:|||||
Db 525 SerLeuMetAlaGluIleuHISProAsnIleValcysLeuLeuGlyAlaValThrGln 544
QY 118 ATCCACCCCGCTGC-----TTCGCC 138
   |||:|||||
Db 545 GluGluProValcysMetLeuPheGluTyrIleAsnGlnGlyAspLeuHISgluPheLeu 564
QY 139 CTGGAGCTGGCGCGCTGACGACCTTCACACACCGCTGCGCAGAGACGACAGATTTC 198
   |||:||||| ||| ||| |||:|||||
Db 565 IleMetArgSerProHISSerAspValGlyCysSerSerAspGluAspGlyThrValLys 564
QY 199 TCCTTTATACCCCTGGGACACATGCTCACCACAAAATAGCCCTACACATCGCCCTGGCG 258
   |||:||||| ||| ||| |||:|||||
Db 585 SerSerLeuAspHISglAspPheLeu-----HISleAlaIleGlnIleAlaIleGly 602
QY 259 CTGGCTTACCTGACACAGAAAACATCATCTTCTGTGACTGTAGTCGAGCGACACATCTCG 318
   |||:||||| ||| ||| |||:|||||
Db 603 MetGluTyrIleuSerSerHISpHePheValHISLysAspLeuAlaAlaArgAsnIleLeu 622
QY 319 GTGTGTCCTTACGCTCAAGACATCATCATCATCATCATCATCATCATCATCATCATCATCG 378
   |||:||||| ||| ||| |||:|||||
Db 623 -----IleGlyGluGlnLeuHISValLysIleSerAspLeuGlyLeuSer 657
QY 379 AGGAGTCATTCATGAGGCGCGCTAGGCGTGAGGCGCATCTCT-----GGC 426
   |||:||||| ||| ||| |||:|||||
Db 638 ArgGluIleTyrSerAlaAspTyrTyrArgValGlnSerLysSerLeuLeuProIleArg 657
QY 427 TACCAGGCCCCAAGATCAGAGCTCCCATGTATATGATAGAGATGATATGTTCTCC 466
   |||:||||| ||| ||| |||:|||||
Db 658 TyrMetProProIleAlaIleMetTyrGlyLysPheSerSerAspSerAlaIleTyrPser 677
QY 487 TATGGAATGGTGTCTACAGATGCTGCTCA-----GGACAGCGCCCTGACATGGCGCACAC 543
   |||:||||| ||| ||| |||:|||||
Db 678 PheGlyValValIleuTyrGlnIlePheSerPheGlyLeuGlnProTyrTyrGlyPheSer 697
QY 544 CACCTCCAGATTCACAGAAAGCTGTCCAGAGGCGATCCGCGGTTCTGGGCGACCGCGAG 603
   |||:||||| ||| ||| |||:|||||
Db 698 AsnGlnGluValIleGlnMetValArgLys-----ArgGlnLeuLeuProCysSerGlu 715
QY 604 GAAGTCAGATTCGCGGACCTGACAGGCGCTCATGATGAGTGTGGACATTAAGCCAGAG 663
   |||:||||| ||| ||| |||:|||||
Db 716 AspCysPro---ProArgMetTyrSerLeuMetThrGlnTyrStrpAsnGluIleProSer 734
QY 664 AAGGAGCGGTGGCGCTGTGGGTGAGGACAGATGAAGGACCGGATTTTGCACACTTC 723
   |||:||||| ||| ||| |||:|||||
Db 735 ArgTyrProArgPheLysAspIleHISValArgLeuArgSer----- 748
QY 724 ATGTATGACTGTGCTGTGGGAGACAGACGCTTCTTCATCCAGGCGCAGAGACTAC 783
   |||:||||| ||| ||| |||:|||||
Db 748 ----- 748
QY 784 ACCGTGTGTTTGGATGCA-----AAAGAGAGACTCCAGG 819
   |||:||||| ||| ||| |||:|||||
Db 749 -----TrrGluGlyLeuSerSerHISThrSerSerThrThrProSerGlyGly 764
QY 820 AACCTACAGGATGGGAGACAGAGAGGCGCTCATGAGGTCGAGAGATGTGTCGCT 879
   |||:||||| ||| ||| |||:|||||
Db 765 AsnAlaThrThrGlnThrHISerLeuSerAlaSerProValSerAsnLeuSerAsnPro 784
QY 880 GCGATGAAGTGAAGCTCCAGCTCCAGGTCCAGAGATCCCTGTGACAGACCGACGAGAC 939
   |||:||||| ||| ||| |||:|||||
Db 785 -----Arg 785
QY 940 CAGAAATCTACATCTACACCCCTCAAGGCGATGCGCCCTTAACACACCCCAACAGGCC 999
   |||:||||| ||| ||| |||:|||||
Db 786 TyrProAsnTyrMetPheProSerGlnGlyLe-----ThrProGln----- 799
QY 1000 TTGGATATCTCAGCTGTGCTGCTCTTGGCGCGTGTGCTGTATTAAGAAATTC 1059
   |||:||||| ||| ||| |||:|||||
Db 800 -----GlyGlnIleAlaGlyPheIleGlyProProIleProGlnAsnGlnArg 815

```

```

QY 1060 TACCTGTCTTAGCGGCGCTCCGCCATGGGCTTGTGGCTGTGTTCCCTGGTGGCGGCC 1119
   |||:||||| ||| ||| |||:|||||
Db 816 PheIleProIleAsnGlyLys---ProIle----- 824
QY 1120 ACCCCAAAGAGACGCTGCTTACCTGTGCTGCTCACACAGACGCAACAGT----- 1168
   |||:||||| ||| ||| |||:|||||
Db 825 -ProProGlyTyrAlaAlaPheProAlaAlaHISLysTyrGlnProThrGlyProProArgVal 844
QY 1169 -----CCAGTTACAGATCGCGGATGAAGACGACGCGAGAACCC 1209
   |||:||||| ||| ||| |||:|||||
Db 844 IleGlnHIScysProProProLysSerArgSerProSerSerAlaSerGlySerThrSer 864
QY 1210 TACC 1213
   |||
Db 864 rThr 865

RESULT 9
US-08-222-616-20
Sequence 20, Application US/08222616
Patent No. 5635177
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,616
FILING DATE: 4-Apr-1994
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00586
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/826935
FILING DATE: 22-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 821P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-222-616-20

Alignment Scores:
Pred. No.: 1,15e-10 Length: 505
Score: 243.50 Matches: 75
Percent Similarity: 50.41% Conservative: 49
Best Local Similarity: 30.49% Mismatches: 77
Query Match: 6.00% Indels: 45

```



```
Db 194 ValValLeuTriPoliMetLeuThrArgGluValProPheLysGlyLeuGluGlyLeuGln 213
QY 553 ATTCCAGAACCTGTCCAGAGGCATCCGCCGTTCTGGGGCAGCCGAGAGAGTGCAG 612
Db 214 ValAlaTriPleu-----ValValGluLysAsnGluArqLeuThr 226
QY 612 TTC-----CGAGACTGCAGCGCCTCATGATGAGTGTGGGACACTAAG 657
Db 627 IleProSerSerCysProArqSerPheAlaGluLeuLeuHisGlnCysTriPoliAlaAsp 246
QY 658 CCAGAGAACGACGACCGCTGTGCGTGTGAGCCAGATGAGAGACCCGACTTTGCC 717
Db 247 AlaLysLysArgProSerPheLysGlnIleIleSerIleLeuGluSerMetSerAsnAsp 266
QY 718 ACCTTCATGTATGAACTGTGC 738
Db 267 ThrSerLeuProAspLysCys 273
```

Search completed: April 15, 2003, 07:22:16
Job time : 40.5 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 15, 2003, 07:08:22 ; Search time 52 Seconds

(Without alignments)
5135.427 Million cell updates/sec

Title: US-09-836-392-8_COPY_22_2205

Perfect score: 4058
Sequence: 1 atgcgagcgactgcgcggc.....ctgcactgcgaagaag 2184

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 6.0 , Xgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 497624

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p model -DEV=xlh
-O=/cgn2_1/USPFO/US09836392/unat_08042003_090330_22925/app_query_fasta_1.2375
-DB=published.Applications_AA -QFMT=fastan -SUFFIX=n2p.rabp -MINMATCH=0.1
-LOOPCL=0 -LOOEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsun62
-TRANS=man40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09836392.ecgn1.1.24.ernat.08042003.090330.22925
-NCPU=6 -ICPU=3 -NO_XIPY -NO_XMAP -LARGEMEMORY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=130 -WARR.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEM_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3822	94.2	728	US-09-836-392-21	Sequence 21, Appl
2	3574.5	88.1	1987	US-10-132-382-6	Sequence 6, Appl
3	3574.5	88.1	2013	US-10-132-382-2	Sequence 2, Appl
4	3574.5	88.1	2014	US-10-132-382-8	Sequence 8, Appl

5	3574.5	88.1	2040	US-10-132-382-4	Sequence 4, Appl
6	337.5	8.3	847	US-10-143-133-2	Sequence 25, App
7	291	7.2	1036	US-09-771-161A-255	Sequence 256, App
8	291	7.2	1036	US-09-771-161A-256	Sequence 2, Appl
9	287	7.1	1036	US-10-014-882-2	Sequence 19, Appl
10	279.5	6.9	1394	US-09-862-027-19	Sequence 43, Appl
11	255.5	6.3	626	US-09-862-027-43	Sequence 2, Appl
12	255.5	6.3	752	US-09-835-081-2	Sequence 48, Appl
13	252	6.2	628	US-09-862-027-48	Sequence 129, App
14	244	6.0	937	US-09-974-298-129	Sequence 6, Appl
15	243.5	6.0	505	US-09-977-260-6	Sequence 6, Appl
16	243.5	6.0	505	US-09-977-261-6	Sequence 6, Appl
17	243.5	6.0	505	US-09-977-262-6	Sequence 20, Appl
18	243.5	6.0	505	US-09-982-610-20	Sequence 5, Appl
19	243	6.0	455	US-09-757-982-5	Sequence 2, Appl
20	242.5	6.0	850	US-09-904-389-2	Sequence 40, Appl
21	242.5	6.0	1601	US-09-862-027-40	Sequence 197, App
22	240.5	5.9	966	US-09-771-161A-197	Sequence 47, Appl
23	238.5	5.9	627	US-09-862-027-47	Sequence 29, Appl
24	237	5.8	425	US-09-828-313-29	Sequence 80, Appl
25	237	5.8	1265	US-09-862-027-80	Sequence 79, Appl
26	236.5	5.8	1308	US-09-862-027-79	Sequence 5, Appl
27	236	5.8	263	US-09-840-704-5	Sequence 81, App
28	234.5	5.8	1734	US-09-862-027-81	Sequence 235, App
29	233	5.7	795	US-10-108-605-235	Sequence 30, Appl
30	232.5	5.7	251	US-08-987-689A-30	Sequence 4, Appl
31	232	5.7	579	US-10-158-895-4	Sequence 15, Appl
32	232	5.7	590	US-10-158-895-15	Sequence 53, Appl
33	227.5	5.6	824	US-09-909-567B-53	Sequence 26, Appl
34	226.5	5.6	912	US-08-281-417-26	Sequence 107, App
35	226.5	5.6	968	US-09-281-417-107	Sequence 7, App
36	225	5.5	1317	US-09-963-896-7	Sequence 2, Appl
37	220	5.4	1167	US-10-274-409-2	Sequence 24, Appl
38	220	5.4	1276	US-09-982-610-24	Sequence 8, Appl
39	219.5	5.4	835	US-09-947-199-8	Sequence 78, Appl
40	219.5	5.4	1237	US-09-862-027-78	Sequence 30, Appl
41	218.5	5.4	656	US-09-862-027-30	Sequence 10, Appl
42	218	5.4	265	US-09-797-039-10	Sequence 4, Appl
43	218	5.4	278	US-09-842-582-4	Sequence 4, Appl
44	218	5.4	278	US-09-934-406-4	Sequence 13, Appl
45	218	5.4	278	US-09-797-039-13	

ALIGNMENTS

RESULT 1
US-09-836-392-21 Application US/09836392
Sequence 21, Appl
Patent No. US20020173458A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
FILE REFERENCE: PTO2001
CURRENT APPLICATION NUMBER: US/09/836,392
CURRENT FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: PCT/US00/28066
PRIOR FILING DATE: 2000-10-11
PRIOR APPLICATION NUMBER: 60/159,542
PRIOR FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: 60/165,914
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: 60/189,027
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 728
TYPE: PRT
ORGANISM: Homo sapiens
US-09-836-392-21
Alignment Scores:

Prod. No.: 8.67e-215 Length: 728
Score: 3822.00 Matches: 728
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.18% Indels: 0
DB: 9 Gaps: 0
US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-836-392-21 (1-728)

QY 1 AGCGAGGACACTGGGGCCACCGATGCGATGAGAGAACTTCTCCGAGTCCGGCGAGG 60
DB 1 MetLeuAaGHisLeuValThrAlaThrAspAlaMetLysAsnPhseSerGluPheArgGlnGlu 20
QY 61 GCCAGATGCTGCAGCGCTGCAGCACCCCTCATGCTGGCGCTCATCGCATCCAGCATC 120
DB 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
QY 121 CACCGGCTTGGTGGCCCTGAGAGCTGGCGCGCTCAGCAGCCTCAACACCGTGTGTC 180
DB 41 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
QY 181 GAGAACGCCAGAGATCTCTTATACCCCTGGGACACATGCTCCACCAAAAATAGCC 240
DB 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 80
QY 241 TACCAAGATCGCTCGGGCTGGCTACCTGACACAGAAAACATCATCTTCTGTGACTG 300
DB 81 TyrGlnIleAlaSerGlyLeuValaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
QY 301 AAGTGGGACACATCTGTGTGTGCTCCCTTACAGCTCAGAGACACATCAATCAAGTA 360
DB 101 LysSerAspAsnIleLeuValaTyrPheSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
QY 361 TCTGACTAGGGATTTCCGAGCAGTCATTCAGAGGGCGCCCTGAGCGCTGAGGCACT 420
DB 121 SerAspTyrGlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnGlyThr 140
QY 421 CCTGGCTACAGGCCCCAGAGATCAGGCGCTCGCATTTGATGATGAGAAGTATGATG 480
DB 141 ProGlyTyrGlnAlaProGlnIleArgProArgIleValTyrAspLeuLysValAspMet 160
QY 481 TTCTGCTAATGAGTGTCTTACAGAGTGTGTGTCAGAGACCGCCCTGACTGGGCGAC 540
DB 161 PheSerTyrGlyMetValaLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHis 180
QY 541 GACCGACTCCAGATTGCCAAGAGCTGTCCAGGGCATCCGCGCTTGGGGCGAGCG 600
DB 181 HisGlnLeuGlnIleLeuLysLysLeuSerLysGlyIleArgProValLeuGlnPro 200
QY 601 GAGGAAGTGCAGTTCGGGCTGCAGGCGCTCATGATGAGTGTGAGACATAAGCA 660
DB 201 GlnGlnValaGlnPheArgTyrGlnGlnAlaLeuMetGlnCysTrpAspThrLysPro 220
QY 661 GAGAAAGCGCCCTGGCCCTGTGCGTGTGAGACCCAGATAGACCCGACTTTGGCCAC 720
DB 221 GlnLysArgProLeuAlaLeuSerValaLysGlnMetLysAspProThrPheAlaThr 240
QY 721 TTCATGTATGAAGTGTGCTGGGAAAGCAGACAGCCTTTCTCAATCCAGGGCGAGG 780
DB 241 PheMetTyrGlnLeuCysGlyLysGlnIleAlaPhePheSerSerIleGlnGlnGlu 260
QY 781 TACACCGTGTGTTTGGATGAGAAAGAGAGTCCAGGAACTACACGCTGTGTAACACA 840
DB 261 TyrThrValValaPheTrpAspGlyLysGlnGlnLysArgAsnTyrThrValaLysnThr 280
QY 841 GAGAAAGGCTCATGAGAGTGCAGAGAGATGTCTGCGCTGGAGATGAAGTGAAGTCCAG 900
DB 281 GlnLysGlyLeuMetGlnValaGlnArgMetCysCysProGlyMetLysValSerCysGln 300
QY 901 CTCGAGGCTCAGAGATCCCTGTGAGACAGCCAGGAGACAAATAATCAATCAATCAAC 960
DB 301 LeuGlnValaGlnArgSerLeuThrAlaThrGlnLysGlnLysIleTyrIleTyrThr 320

QY 961 CTCAGAGCATGTGCCCTTAAACACACCCCAACAGCCCTTGATACTCCAGTGTCTTC 1020
DB 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
QY 1021 ACCTGCTTCTTGGCGCGCTGTGTATTAAGAAATCTTACTGCTTACGCGGCGCTC 1080
DB 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
QY 1081 GCCGATGGGCTGTGGCTGTGTTTCCCGTGGTGGCGGGGACCCCAAGAGACAGTCTCC 1140
DB 361 AlaAspGlyLeuValaAlaValaPheProValaValaArgGlyThrProLysAspSerCysSer 380
QY 1141 TACCTGTGCTCACACAGCCAAAGGCTCAAGTTCCAGCATCGCGGATGAAGACGACAG 1200
DB 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspLysArg 400
QY 1201 CAGAACCCCTACCCAGTGAAGGCGCATGAGGTGTCAACAGCGCTGTAGGTGTGTAC 1260
DB 401 GlnAsnProTyrProValLysAlaMetGluValaValaAsnSerGlySerGluValaTyrPyr 420
QY 1261 AGCAATGGCGCGGCTGCTGTGATGCACTGTGCTCCCTGAGATGTCAGCGGCGCTG 1320
DB 421 SerAsnGlyProGlyLeuLeuValaIleAspCysAlaSerLeuGlnLecysArgArgLeu 440
QY 1321 GAGCCCTACATGGCCCTCCATGGTTACATGCTGCTGTGAGCTGTAGGCGCAGAGG 1380
DB 441 GluProTyrMetAlaProSerMetValaHisSerValaCysSerSerGlnLysArgGly 460
QY 1381 GAGAGAGTGTGTGCTGCTGTGATGACAGAGCCAACTCTTGATGTATACCATCCACC 1440
DB 461 GlnGlnValaValaTyrCysLeuAspAspLysAlaAsnSerLeuValaMetTyrHisSerThr 480
QY 1441 ACCTACAGCTGTGTGCGCGGCTGCTGCGGGGCGCCAGCGCCCTGAGGACATGTTT 1500
DB 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValaProSerProLeuArgAspMetPhe 500
QY 1501 CCCGTGCGGCGCTTGGACAGAGGAAACCCCGGAGCAGCAGCAGCCCAAGAGTG 1560
DB 501 ProValaArgProLeuAspThrGlnLysProProAlaAlaSerHisThrAlaAsnProLysVal 520
QY 1561 CCTGAGGGGAGCTCCATGCGCGGAGCTGAGACATCATGTACAGTGAAGAGCTGGGCGAG 1620
DB 521 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGlnLeuGlnThrGln 540
QY 1621 ATCTGATCCACAGGATATCATCTGACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 541 IleLeuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSer 560
QY 1681 CCACCCGCGAGGCTGCAGAGTCCCGCTCAAGCGCTCCCGAGCTCCCGAGCAAGTCTTCC 1740
DB 561 ProProAlaArgAlaAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
QY 1741 AGTGTGCTTTTCCACAGAGTGCAGAGGATCTGACAGATGCTATAGCCCGGCTGCTCC 1800
DB 581 SerValaProPheSerThrAspCysGlnAspSerLeuAspMetLeuHisThrProGlyAlaAla 600
QY 1801 TCCGACAGTGTGACATGAGCTGACCCCGCATGAGCGGGGAGAGCTTGAAGCAGAGCTG 1860
DB 601 SerAspArgSerGlnHisSerPheLeuThrProMetAspGlyGlnThrPheSerGlnHisLeu 620
QY 1861 GAGCCGCTGAAGATCCCTGCGCTGCAGAGACCTCATTTGGTCCCGGCGGCTGTGAGAT 1920
DB 621 GlnAlaValaLysIleLeuAlaValaLysArgPheLeuIleTyrValProAlaArgGlyLysP 640
QY 1921 GTTATGCTATTTGGCTGAGAGAGGATTTCTGAAGCCCAAGCGGGCGCGATCTTCCCTC 1980
DB 641 ValIleValaIleGlyLeuGlnLysAspSerGlnAlaGlnArgGlyArgValaIleAlaVal 660
QY 1981 TTAAGCCCGAGAGCTACCTCCGATGGGAGTGGGTGAGATGCTGCGCTGTGAGCAAG 2040
DB 661 LeuLysAlaArgGlnLeuThrProHisGlyValaLeuValaAspAlaValaValaLys 680
QY 2041 GACACTGTTGTGTGACACTTTGAAAATGAAAACAGAGTGTGCTGCGCTGAGAG 2100

Db 681 AsprValValcysThrPheGluAsnGluAsnThrGluTrpCysLeuAlaValTrpArg 700
QY 2101 GGCTGGGGGGCCAGGAGATTGACATTTTACACGCTCTACAGAGACTGGCGCGCTG 2160
Db 701 GlyTrpGlyAlaAlaArgLysPheAspIlePheTyrGlnSerTyrGlnGluLeuGlyArgLeu 720
QY 2161 GAGGCTTCACATCCGACAGAGAGG 2184
Db 721 GluAlaCysThrArgLysArgArg 728
RESULT 2
US-10-132-382-6
: Sequence 6, Application US/10132382
: Publication No. US20030045699A1
: GENERAL INFORMATION:
: APPLICANT: WEISS, BERTRAM
: TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
: FILE REFERENCE: SCH-1811
: CURRENT APPLICATION NUMBER: US/10/132,382
: CURRENT FILING DATE: 2002-04-26
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 1987
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-132-382-6

Alignment Scores:
Pred. NO.: 2.43e-200 Length: 1987
Score: 3574.50 Matches: 687
Percent Similarity: 94.50% Conservative: 1
Best Local Similarity: 94.50% Mismatches: 1
Query Match: 88.09% Indels: 39
DB: 9 Gaps: 1

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-132-382-6 (1-1987)

QY 1 AAGCTGAGGACCTGGGGGCGACCGATGACATGAGAACTTCCGAGTCCGGCAGAG 60
Db 1300 MetLeuAlaThrIleuAlaThrAlaThrAlaMetLysAsnPheSerGluPheArgGlnGlu 1319
QY 61 GCCAGCATCTGCAGCGCTGCAGCACCCTGATCGTGGCGCTCATCGCATCAGCATC 120
Db 1320 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 1339
QY 121 CACCGCTCTGCTGGCCCTGAGCTCGCGCGCTCAGCAGCCTCAACACGCTGCTCC 180
Db 1340 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 1359
QY 181 GAGAACGCGAGATCTCTCTTATACCCCTGGGACACATGCTCCCAAAAATAGCC 240
Db 1360 GluAsnAlaArgSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 1379
QY 241 TACCAATGCGCTCGGCGCTGAGCTACATGACACAAAAACATCATCTTCTGACCTG 300
Db 1380 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 1399
QY 301 AAGTCGACACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 1400 LysSerAspAsnIleLeuValTrpSerLeuAspValLysGlnHisIleAsnIleLysLeu 1419
QY 361 TCGACTAGGGGATTTCGAGGACATTCATTCATGAGGCGGCTAGCGGTGAGGCGACT 420
Db 1420 SerAspTyrGlyIleSerArgInsSerPheHisGlnGlyAlaLeuGlnValGlnGlyThr 1439
QY 421 CCGTGTACACAGGCGCCAGATCAGAGCTCGCATTTGATATGATGAGAAGTAGATATG 480
Db 1440 ProGlyTyrGlnIleAlaProGlnIleArgProArgIleValTyrAspGlnLysValAspMet 1459
QY 481 TTCTCTATGAAATGGTCTCTACAGATGCTCTCAGAGACGCGCTGCTGAGGCGAC 540

Db 1460 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHis 1479
QY 541 CACGAGCTCCAGATTGCCAAGAGCTGTCCAGAGGATCCCGCTTCTGCGGCGAGCCG 600
Db 1480 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnArgLeu 1499
QY 601 GAGAGATGCAATTCGCGGCGACTGACGCGCTCATGATGAGTGTGCGGACACTAAGCCA 660
Db 1500 GluGlnValAlaGlnPheArgArgLeuGlnAlaLeuMetGlnCysTrpAspThrLysPro 1519
QY 661 GAGAGCGACCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 1520 GluLysArgProLeuAlaLeuSerValAlaSerGlnMetLysAspProThrPheAlaThr 1539
QY 721 TTCATGATGAATGCTGTGTGGAGACAGACAGCCTTCTCTATCCAGGCGCAGAG 780
Db 1540 PheMetTyrGlnLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlnGln 1559
QY 781 TACACCGTGTGTTTGGATGAGAAAGAGAGTCCAGGAATACAGCGGTGTGAACA 840
Db 1560 TyrThrValAlaPheTrpAspGlyLysGlnGlnSerArgAsnTyrThrValAlaAsnThr 1579
QY 841 GAGAGCGCTCATGAGAGTGCAGAGATGTCGCTCCCTGGGATGAGTGAAGTGCAG 900
Db 1580 GluLysGlyLeuMetClnValAlaGlnArgMetCysCysProGlyMetLysValSerCysGln 1599
QY 901 CTCAGAGTCCAGAGATCCCTGTGTGACACGACCAAGGACCAAGAAATCTACATTCAC 960
Db 1600 LeuGlnValGlnArgSerLeuThrPheThrAlaThrGln 1611
QY 961 CTCAAGGCGATGCGCCCTTAACACACCCCAACAGGCTTGATATCTCCAGCTGTCTC 1020
Db 1611 ----- 1611
QY 1021 ACCGTCTTCTGGCGCGCTGTATTAAGAAATTCCTACCTGCTTACGCGGCGCTC 1080
Db 1612 ----- 1612
QY 1081 GCCGATGGGCTTGTGTGTGTGTGTCCCGTGTGCGGGGACCCCAAGAGACGTGCTCC 1140
Db 1621 AlaAspGlyLeuValAlaValaPheProValValaArgGlyThrProLysAspCysSer 1640
QY 1141 TACCTGTGCTCACACACAGCCCAAGAGTCCAAAGTTCAGATGCGGATGAACAGCAG 1200
Db 1641 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlnAspAlaArg 1660
QY 1201 CAGAACCCCTACCCAGTGAAGGCGCATGAGGTGTCGAACAGGCGCTGAGGCTGCTG 1260
Db 1661 GluAsnProTyrProValLysAlaMetGlnValValaAsnSerGlySerGlnValTrpTyr 1680
QY 1261 AGCAATGGCGCGGCGCTCTGTGATCGACTGTGCTCCCTGGAGATCTGACAGCGGCTG 1320
Db 1681 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgArgLeu 1700
QY 1321 GAGCCCTACATGGCCCTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1701 GluProTyrMetAlaProSerMetValThrSerValValaCysSerSerGlnGlyArgGly 1720
QY 1381 GAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Db 1721 GluGlnValValaTrpCysLeuAspAspLysAlaAsnSerLeuValMetLysHisSerThr 1740
QY 1441 ACCATCAGCTGTGTGCTCCCGGATCTTCTGCGGGGCTCCAGCGCCCTCAGGAGCATGTT 1500
Db 1741 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValaProSerProLeuArgAspMetPhe 1760
QY 1501 CCGGTGCGGCGCTTGGACACGGAACCCCGGACGCAAGCAGCGGCAACCAAGGTG 1560
Db 1761 ProValArgProLeuAspPheArgLysProProAlaAlaSerHisThrAlaAsnProLysVal 1780
QY 1561 CCGTGGAGGAGATCCATCGCGGACGAGCATCATGATGATGATGATGATGATGATGATG 1620
Db 1781 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGlnGlnGlnGlnGln 1800

```

QY 1621 ATCTGATTCACAGAGATCACTACTGCTGCTCCATGCTCTACTGCTCATCC 1680
Db 1801 IIELEUILLIHSGLINLUERLEUTHRASPTRYCYSSERMESSEYRSEYRSEYR 1820
QY 1681 CCACCCCGCCAGGCTGCGAGGTCCCGCTCAAGCTCCCGAGCTCCCGACAGTTCTTC 1740
Db 1821 PROFIARGLINLAHLARSERPROSERSEYRSEYRSEYRSEYRSEYRSEYR 1840
QY 1741 AGTGTGCTTTTCTCCACCGACGCTGCGAGACTAGACATGCTACATACGCGCGTCTGCC 1800
Db 1841 SERVALPROPHESERTHRSPCYSGIUSASERASPMELIENHISTHRPROGLIYALAIA 1860
QY 1801 TCCGAGAGCTGAGCATACCTGACCCCATGAGAGGGGAGACCTTGACCGACAGCTG 1860
Db 1861 SERASPARSERGLIHISASPLEUTHRPMELASPLGLIUTHRPHESERGLIHISLEU 1880
QY 1861 CAGGCGGTGAGATCTCCCGCTCAGAGACCTCATTTGGGTCCCGACGCGGTGAGAT 1920
Db 1881 GINLAVALIYSLILEUHLAVALARGASPLEUILLIETRYVALPROHARGGLIYASAP 1900
QY 1921 GTTATGCTATGCTGCTGAGAGAGATTTGTGAAGCCCGGCGGCGAGTCAATGCCCTC 1980
Db 1901 VALIIEVALIILEUHLIYSLIYUSASPSERGLYALAGIHLARGIYARGVALIIEALIA 1920
QY 1981 TTAAGAGCCGAGAGAGCTGACCTCGCATGGGTGCTGCTGATGCTCGGAGTGGCAAG 2040
Db 1921 LEULYALARGGLIUEUTHRPROHISGLIYALILEUVALIASPILALIAVALIALYLS 1940
QY 2041 GACACTGTTGTGTGACCTTTGAAAAATGAACACAGAGTGTGCTGCGCGCTGGAGG 2100
Db 1941 ASPTHLVALIYALICYSRHPHEGLIUSNGIUSANTHGLIUTRIPCYSLEUHLAVALITR 1960
QY 2101 GGCTGGGCGCGAGAGAGTTCGACATTTCTACCATCTCTAGAGAGAGTGGCGCGCTG 2160
Db 1961 GLYTPRGLYALARGGLIUEUTHRPHIENHISPIEHPHYRGLINSEYRGLIUEUGLYARG 1980
QY 2161 GAGGCTTGACCTGCGAAGAGA 2181
Db 1981 GIUALICYETHRARGLYSARG 1987

RESULT 3
US-10-132-382-2
; Sequence 2, Application US/10132382
; Publication No. US2003004569A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2013
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-2

Alignment Scores:
Pred. No.: 2,44e-200 Length: 2013
Score: 3574.50 Matches: 687
Percent Similarity: 94.50% Conservative: 0
Best Local Similarity: 94.50% Mismatches: 1
Query Match: 88.09% Indels: 39
DB: 9 Gaps: 1

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-132-382-2 (1-2013)
QY 1 ANGCGAGGACACCTGGGCGCCAGATGCGCATGAGAGAAATTTCTCGAGTTCGCGCAGAG 60
Db 1336 NCTLEUARGHISLEUARGALATHRASPALAMELYASNPHESEYRGLIUEHARGGLIN 1345

```

```

QY 61 GCGAGATGCTGACGCGCTGAGAGACCCCTGATGCTGGCGCTCATGCGCATC 120
Db 1346 ALASERMELIENHISLALEUHLINISPROCYSLIEVALIALEUHLILEYLI 1365
QY 121 CACCGCTCTGCTGCGCTGAGAGTCCGCGCTCAGAGCTCAAGCTCAACACCTGCTGC 180
Db 1366 HISPROLEUCYRPHALALEUHLIUEUHLAIPROLEUSERSEYRLEUHLNTHRVALEUSER 1385
QY 181 GAGAAAGCCAGAGATTTCTTTATACCCCTGGAGACACATGCTCAACCAAAAAATAGCC 240
Db 1386 GIUASNALARGASPSERSEYRHEILEPROLEUGIYNHISMETLEUTHCINLYSILEA 1405
QY 241 TACGAGATGCGCTGCGCGCTGAGCTACCTGACAGCAAGAAAATATATCTTCTGATCCTG 300
Db 1406 TYRGLIIEALISERGLIUEUHLATRYLEUHLISLYSASNIILEIHPHECYASPLEU 1425
QY 301 AAGTGGAGACATTTCTGCTGTGCTCCCTTGAGCTCAAGAGACACATCAATCAAGCTA 360
Db 1426 LYSERASPARSNIILEUVALITRPSERLEUAPVALIYSGIHLIIEASNIIIEYSLI 1445
QY 361 TCTGACTACGGGATTTCCAGGCAATTCATTCATGAGGCGCGCTGAGGCGGCACT 420
Db 1446 SERASPTRYGLIYLESERARGIINSEYRPHENISGLIUGIYALILEUGIYALIGI 1465
QY 421 CCTGGCTACAGGCGCCAGAGATCAGGCGCTGCTGATGATGATGAGAGGATAGATATG 480
Db 1466 PROGLIYRGLIHALPROGLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLI 1485
QY 481 TTCTCTATGAGATGCTGCTACAGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 1486 PHESERTYRGLIYMEVALIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLI 1505
QY 541 CACGAGCTTCCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 1506 HISSGLINLEUHLIIEALIALYSLIYUSLEUSERLYSGIYIIEARGPROVALILEUHLI 1525
QY 601 GAGAGAGTCACTGCTGCGGAGCTGCGAGGCTCATGATGAGAGTGGGAGACTAAGCA 660
Db 1526 GIUGIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLI 1545
QY 661 GAGAAAGCAGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 1546 GIULYSARGPROLEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLI 1565
QY 721 TTCATGATGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 1566 PHMETIYRGLIUEUCYSCYSGIYSLIYSGIINTHRALAPHEPHESEYRGLINLI 1585
QY 781 TACACCGTGTGCTTTGGATGAGAAAGAGAGTCCAGAACTACAGCTGCTGAACACA 840
Db 1586 TYRTHRVALIYALPHETRPSPOLYSGIUGIUSERARGASNTTYRTHVALIALSNTHR 1605
QY 841 GAGAAAGCCTCATGAGAGTGCAGAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 1606 GIULYSGLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLI 1625
QY 901 CTCGAGGTCCAGAGATCCCTGTGAGACAGCCAGCGAGACAGAAATCTACATCTACACC 960
Db 1626 LEUGIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLI 1637
QY 961 CTCAGAGCATGTGCCCTTAACACACACCCCAAGGCGCTTGATATCCAGGTGCTGCTC 1020
Db 1637 ----- 1637
QY 1021 ACCTGCTTCTGCGCGCTGCTGCTTATTAAGAAATCTCTACCTGCTGCTGCTGCTGCTG 1080
Db 1638 -----AANSERTYRLEUVALIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLI 1646
QY 1081 GCGGATGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 1647 ALAASPTGLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLIUEUHLI 1666
QY 1141 TACCTGTGCTCACACACAGCCCAAGGTTCAGATTCACATGCGGAGTGAAGACGACGAG 1200

```

```

Db 1667 TyrlucysserlnstrlalaasnargserlyspheSerllealaaspgluaSpalaarg 1686
QY 1201 CAGAACCCCAACCCAGAGAGCCCATGAGGTGGTCAACAGCGGCTGTGAGGCTGTAC 1280
Db 1687 GluaSnProtyrProvalylsalameGluvalalasnserGlyserGluvaltlrpyr 1706
QY 1261 AGCAATGGCGCGGCTCTGTGTCATGACTGTGCTCCCTCGAGATCGAGCGGCTG 1320
Db 1707 SerlnsnglyProglyleuLeuVallleaspcyslaSerleuGluilleGysargArgleu 1726
QY 1321 GAGCCCTACATGAGCCCTCCATGCTTACGTCACTGTGTGACCTGTGAGGAGAGG 1380
Db 1727 GluProtyrMetalaProSerMetValtlrSerValValCysSerSerGluGlyArgGly 1746
QY 1381 GAGAGGTGCTGTGCTGTGATGACAGAGCCCAACTCCTGGGAGTGTACCATCCACC 1440
Db 1747 GluGluValValtlrPcyslaasnspasylsalasnSerleuValMetTyrHisSerThr 1766
QY 1441 ACCTACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1500
Db 1767 ThrTyrGlnleuCySlaalaArgTyrPheCysGlyValProSerProleuArgAspMetPhe 1786
QY 1501 CCCGTGGCGGCTGTGACAGCGAAACCCCGGACGACCCACAGCGGCGCAAGGTG 1560
Db 1787 ProValArgProleuAspThrGluProProAlaalaSerlnstrlalaasnProlysal 1806
QY 1561 CCTACGGGAGCTGCATCGGCGAGGTGAGCATGATGATGATGAGGAGGAGGCGGAG 1620
Db 1807 ProGluGlyAspSerllealaSpasylSerlleMetTyrSerGluGluGlyThrln 1826
QY 1621 ATCTGATCAACAGAAATCACTACTACTACTCTCTCTCTCTCTCTCTCTCTCTCT 1680
Db 1827 lleuIlelnsnglnGlnSerleuThrAspTyrCysSerMetSerSerTyrSerSerSer 1846
QY 1681 CCACCCCGGAGGCTGCGAGGTCCCTCAAGCTCCCGGAGCTCCCGGAGGAGTCTTC 1740
Db 1847 ProProArgGlnAlaalaArgSerProSerSerleuProSerProAlaSerSerSer 1866
QY 1741 AGTGCGCTTCTCCACGAGCGAGGAGAGCTCAGACATGATGATGATGATGATGATG 1800
Db 1867 SerValProPheSerThrAspCysGlnspSerAspMetleuHisThrProGlyAlaala 1886
QY 1801 TCCGACAGGTGTGACATGACCTGACCCCATGACGAGGAGAGACCTTCAGCCAGCCTG 1860
Db 1887 SerAspArgSerGlnHisAspSerThrProMetAspGlyGluThrPheSerGlnHisleu 1906
QY 1861 CAGCCCGGAGAGATCTGCGGCTCAGACATCTCATTTGGGTCCCGAGCGGCTGAGAT 1920
Db 1907 GlnAlaVallylleuAlaValalrAspSerleuIleTyrValProArgArgGlyGlyAsp 1926
QY 1921 GTTATCGCATGGGCTGAGAGAGATTCGAAAGCCAGCGGCGGCGGAGTCATTCGCTC 1980
Db 1927 ValIleValIleGlyleuGlnlyspSerGlyAlaGlnAlaGlyArgValIleAlaVal 1946
QY 1981 TTAAGCCCGAGAGCTGACTCCGATGGGCTGTGTGATGCTGCGGTGTGCGCAAG 2040
Db 1947 leuylsalalargGlnleuThrProHisGlyValleuValAspAlaAlaValValalals 1966
QY 2041 GACACTGTGTGTGCACTTTGAAAAATGAAAAACAGAGTGGTGGCTGGCTGTGAGG 2100
Db 1967 AspThrValValCysThrPheGlnsnglnuasnThrGluTyrPcysleuAlaValtlrParg 1986
QY 2101 GCGTGGGCGGAGGAGAGTTCGATTTTCTACAGTCTTACAGAGAGTGGGCGGCTG 2160
Db 1987 GlyTyrPglAlaAlargGlnuasnspIlePheTyrGlnSerTyrGlnGluIleuGlyArgleu 2006
QY 2161 GAGCTGTGACTTCGCAAGAGA 2181
Db 2007 GluAlaCysThrArgGlyAsarg 2013

```

```

; Sequence 8, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: MEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2014
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-132-382-8

Alignment Scores:
Pred. No.: 2,44e-200 Length: 2014
Score: 3574.50 Matches: 687
Percent Similarity: 94.50% Conservative: 0
Best Local Similarity: 94.50% Mismatches: 1
Query Match: 88.09% Indels: 39
DB: Gaps: 1

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-132-382-8 (1-2014)
QY 1 ATCTGAGGCACTGTGCGGCGCACCGATGCAATGAAGACTTCTCGAGTTCGGCGAGAG 60
Db 1327 MetleuArgHisleuArgAlaThrAspAlaMetlysaSnPheSerGlnPheArgGlnGlu 1346
QY 61 GCCAGACTGTGACAGCGGCTGACAGACCCCTGCATCGTGGGCTGCATGGCATGACATC 120
Db 1347 AlaSerMetleuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 1366
QY 121 CACCGGCTGTGCTGCGGCTGAGAGTCCGCGGCTCAGACAGCTCAACAGCGGCTGTCC 180
Db 1367 HisProleuCySPhleAlaLeuGlnleuAlaProleuSerSerleuAsnThrValleuSer 1386
QY 181 GAGAACGCCAGAGATTCCTTATACCCCTGGGACACATGCTCACCCAAAATAAGCC 240
Db 1387 GluAlaAlaArgAspSerSerPheIleProleuGlyHisMetleuThrGlnlySleAla 1406
QY 241 TACCAATGCGCTGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 1407 TyrGlnIleAlaSerGlyleuAlaTyrleuHisLysLysAsnIleIlePheCysAspLeu 1426
QY 301 AAGTGGCAACATCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 1427 LysSerAspAsnIleleuValtlrPserleuAspVallylGlnHisIleasnIleLysleu 1446
QY 361 TCTGACTAGGGGATTTTCAGAGCATTCATTCATGAGGCGGCGGCTGAGGCGGCACT 420
Db 1447 SerAspTyrGlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnGlyThr 1466
QY 421 CCTGGCTACAGAGGCGGCGGAGATCAAGGCTGCGATTTGATATGATGAGAAAGTAAATG 480
Db 1467 ProGlyTyrGlnAlaProGlnIleArgProArgIleValTyrAspGlnLysValaspMet 1486
QY 481 TTCCTATGGAAGTGTGCTTACAGAGTTCGCTGACAGGACAGCGGCTGCACTGGGCGAC 540
Db 1487 PheSerTyrGlyMetValleuTyrGlnleuLeuSerGlyGlnThrProAlaLeuGlnHis 1506
QY 541 CACAGCTCCAGATTCGCAAGAGCTGTCCAGAGGAGATCCGCGGCTTCGGGAGAGCG 600
Db 1507 HisGlnleuGlnIleAlaLysLysleuSerLysGlyIleArgProValleuGlnPro 1526
QY 601 GAGGAAGTGCATCTCGGCGACTGCAAGGCGCTCATGATGATGAGTGTGCGACACTAAGCA 660
Db 1527 GluGluValGlnPheArgArgleuGlnAlaMetMetGlnCysTyrAspThrLysPro 1546
QY 661 GAGAGCGAGCGGCTGCGGCTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Db 1547 GluLysArgProleuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThr 1566

```

QY	1801	TCCACAGGCTGAGCATGACTGACCCCATGACGGGAGACCTTCAGCCAGCACTG	1861																								
Db	1088	SerisprgsergelunhisasplethnrPrometAspelygluthPheserolnhsleu	1907																								
QY	1861	CAGCCGTGAAGATCTCGCCGTGACAGACCTCATTTGGGTCGCCAGCGCGGTGAAGAT	1920																								
Db	1908	GlnAlaVallylsIleLeuAlaValAlaArgAspleuIleTrpValProArgArgelyllysp	1927																								
QY	1921	GTTATGCTCAATTGGCCCGGAGAGAAATTTCGAAGCCCGGGCGGAGTCAATTGCGCTC	1980																								
Db	1928	ValleValleIleeglyLeuuglyLysAspserGlyAlaGlnArgelylArGyAlleAlaVal	1947																								
QY	1981	TTAAAGCCCGAGAGACTGACTCCGCATGGGGTCTGGTGGATGCTGCCTGTGGCAAG	2040																								
Db	1948	LeuylsAlaArgGlnLeuThrProHISglyValLeuValAspAlaAlaValAlaLys	1967																								
QY	2041	GACACTTTGTGTGACACTTTGAAATGAAACACAGAGATGCTGCGCTGGCTGAGG	2100																								
Db	1968	AspThrValValCysThrPrpGlnAsnGlnAsnThpGlnTrpCysLeuAlaValTrpArg	1987																								
QY	2101	GGCTGGGGGGCCAGGAGGATTCGACATTTTCTACAGACTCTTACGAGAGAGCTGGCCGCTG	2166																								
Db	1988	GlyTrpIlyAlaArgGlnPrpAspIlePheThyGlnserTyGlnGlnLeuGlnlyArgLeu	2007																								
QY	2161	GAGCTTGCACTGGCAAGAGA 2181																									
Db	2008	GlnAlaCysThrArGlyLysArg 2014																									
<p>RESULT 5</p> <p>US-10-132-382-4</p> <p>Sequence 4, Application US/10132382</p> <p>Publication No. US2003004369A1</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: WEISS, BERTRAM</p> <p>TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIS</p> <p>FILE REFERENCE: SCH-1811</p> <p>CURRENT APPLICATION NUMBER: US/10/132,382</p> <p>CURRENT FILING DATE: 2002-04-26</p> <p>NUMBER OF SEQ ID NOS: 26</p> <p>SOFTWARE: PatentIn Ver. 2.1</p> <p>SEQ ID NO 4</p> <p>LENGTH: 2040</p> <p>TYPE: PRT</p> <p>ORGANISM: Homo sapIens</p> <p>US-10-132-382-4</p>																											
<p>Alignment Scores:</p> <table border="0"> <tr> <td>Pred. No.:</td><td>2,44e-200</td><td>Length:</td><td>2040</td></tr> <tr> <td>Score:</td><td>3574.50</td><td>Matches:</td><td>687</td></tr> <tr> <td>Percent Similarity:</td><td>94.50%</td><td>Conservative:</td><td>0</td></tr> <tr> <td>Best local Similarity:</td><td>94.50%</td><td>Mismatches:</td><td>1</td></tr> <tr> <td>Query Match:</td><td>88.09%</td><td>Indels:</td><td>39</td></tr> <tr> <td>DB:</td><td>9</td><td>Gaps:</td><td>1</td></tr> </table>				Pred. No.:	2,44e-200	Length:	2040	Score:	3574.50	Matches:	687	Percent Similarity:	94.50%	Conservative:	0	Best local Similarity:	94.50%	Mismatches:	1	Query Match:	88.09%	Indels:	39	DB:	9	Gaps:	1
Pred. No.:	2,44e-200	Length:	2040																								
Score:	3574.50	Matches:	687																								
Percent Similarity:	94.50%	Conservative:	0																								
Best local Similarity:	94.50%	Mismatches:	1																								
Query Match:	88.09%	Indels:	39																								
DB:	9	Gaps:	1																								
<p>US-09-836-392-8.COPY_22_2205 (1-2184) x US-10-132-382-4 (1-2040)</p>																											
QY	1	ATGTGTAGGAGACCTGGCGGCGCACCGATGCCATGAGAACTTCCGAGTTCGGGAGAGG	60																								
Db	1353	MetLeuArgHisLeuAlaGlnAlaThrAspAlaMetLysAsnPheserGlnPrpArgGlnGln	1374																								
QY	61	GCCAGCATGCTGCACGGCTGCAGACACCCCTGCATCGTGGCGCTCATCGCATCGCATC	120																								
Db	1373	AlaserMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleIlyleSerIle	1399																								
QY	121	CACCGCTGTGCTTCGGCCCTGGAGCTCGGGCGCGCTAGCAGGCTCAACACCGTGTCTG	180																								
Db	1393	HisProLeuLysPrpAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer	1412																								
QY	181	GAGAACCCAGAGATTTCTCTTATACCCCTGGGACATGCTGACACCAAAATAATAGCC	240																								
Db	1413	GlnAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla	1433																								
QY	241	TACCAGATGGCTCGGGCTGGGCTTACTGTGCACAGAAAACATCATCTTGTGCACCTG	300																								

Db 1433 TyrGlnIleAlaSerGlyLeuAlaLeuHisLysLysAlaIleIlePheGlyAspLeu 1452
QY 301 AAGTGGCAACATTTGGTGGTGGTCCCTTACGTCAGAGACATCAACATCAAGCTA 360
Db 1453 LysSerAspAlaIleLeuValThrPheLeuAspValLysGlnHisIleAsnIleLysLeu 1472
QY 361 TCTGCTACGGGATTTCCAGGACATTCATGAGGCGCCCTGAGGGGTGGAGGCGACT 420
Db 1473 SerAspTyrGlyLeuSerArgIleSerPheHisGlnGlyAlaLeuGlyValGlnGlyThr 1492
QY 421 CTTGGCTACAGCCCGCCAGAGATCAGGCTCGCATTTATATGATGAGAGAGTATATG 480
Db 1493 ProGlyTyrGlnAlaProGlyLeuArgProArgIleValTyrAspGlnLysAlaSerMet 1512
QY 481 TTCTCTATGAGATGGTGTCTACAGATTTGTCTGACAGACAGCCCTGACCTGGCCAC 540
Db 1513 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlyHis 1532
QY 541 CACGAGCTCCAGATTCGCAAGAGCTGTCAGGCGCATCCGCGGTTGGGGCAGCCG 600
Db 1533 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProAlaLeuGlyGlnPro 1552
QY 601 GAGAGAGTGCATGCTCCGCGAGCTGACAGGCTCATGATGAGTGTGGAGCACTAAGCCA 660
Db 1553 GlnGlnValGlnPheArgArgLeuGlnAlaLeuMetGlnCysThrAspThrLysPro 1572
QY 661 GAGAGAGCAGCCCTGGCCCTGTGGTGTGTGAGACCAAGATAGACCCGACTTTGGCACC 720
Db 1573 GlnLysArgProLeuAlaLeuSerValLysGlnMetLysAspProThrPheAlaThr 1592
QY 721 TTCATGTATGAACTGTGCTGGGAGAGACAGACGCTTTCTCATCCAGGCGCAGAG 780
Db 1593 PheMetTyrGlnLeuLysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGly 1612
QY 781 TACACCGTGTGTTTGGATGAGAAAGAGAGTCCAGAGACTACAGGCTGTGAACACA 840
Db 1613 TyrThrValValPheThrAspGlyLysGlnGlnLysArgAsnThrValAlaAsnThr 1632
QY 841 GAGAGAGCCTCATGAGAGTGCAGAGAGTGTGTGCTGGGATGAAGTGAAGTGCAG 900
Db 1633 GlnLysGlyLeuMetGlnValGlnArgMetCysArgProGlyMetLysValSerCysGln 1652
QY 901 CTCAGAGTCCAGAGATCCGTGTGAGCAGCAGCAGGAGCAGAAATTCATCACTACAC 960
Db 1653 LeuGlnValGlnArgSerLeuThrPheAlaThrGln 1664
QY 961 CTCAGAGGATGTGCCCTTAAACACACCCCAAGCCTTGATCTCCAGCTGTGCTC 1020
Db 1664 1664
QY 1021 ACCTGCTTCTTGGCCGCTGCTTATTAAGAAATTCCTACTGCTTACGGGCGCTC 1080
Db 1665 1665
QY 1081 GCGGATGGGCTTGGCTGTGTTCCCGTGGTGGGGGCGCCCAAGGACAGCTGCTC 1140
Db 1674 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 1693
QY 1141 TACCTGTGCTCACACAGCAACAGAGTCCAGTCAAGTCCGCGGATGAAGAGCAGG 1200
Db 1694 TyrLeuLysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlnAspAlaArg 1713
QY 1201 CAGAACCCCTACCCAGTGAAGCCATGAGAGTGTCAACAGCGGCTGTGAGTGTGCTAC 1260
Db 1714 GlnAsnProTyrProValLysAlaMetGlnValValAlaAsnSerGlySerGlnValAlaTrpTyr 1733
QY 1261 AGCAATGGGGCGGCGCTTGTCAATGAGTGTGCTGCTGAGAGCTCAGAGCGGCTG 1320
Db 1734 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnLysArgAlaGln 1753
QY 1321 GAGCCATACATGGCGCCCTCATGCTTACGTCAGTCTGTGACGCTGAGGCGAGAGG 1380

Db 1754 GlnProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgGly 1773
QY 1381 GAGGAGCTGTGTGGCTGGATGACAAAGCCAACTCTGTGATGTATACCTCCACC 1440
Db 1774 GlnGlnValValTyrCysLeuAspAspLysAlaAsnSerIleValMetTyrHisSerThr 1793
QY 1441 ACCTACAGCTGTGTGCGGTACTTTCGCGGGTCCCGACCCCTCAGGAGACATGTT 1500
Db 1794 ThrGlnLeuGlyAlaArgTyrPheCysGlyValProSerProLeuAlaArgPhe 1813
QY 1501 CCGCTGGCGCCCTTGGACAGCAAGCCCGCGCAGCAGCCACAGCCCAACCAAGGTG 1560
Db 1814 ProValArgProLeuAspThrGlnProProAlaAlaSerHisThrAlaAsnProLysVal 1833
QY 1561 CCGTGGGGGAGCTCCATCGCGAGCTGAGCATGATGTACAGTGAAGAGAGCTGGAGCAG 1620
Db 1834 ProGlnLysAspSerIleAlaAspValSerIleMetTyrSerGlnGlnLeuGlyThrGln 1853
QY 1621 ATCTGATCCACCAAGAAATCACTACATGACTACTGCTCCATGCTCTCTATCC 1680
Db 1854 IleLeuIleHisGlnGlnLysSerLeuThrAspTyrCysSerMetSerTyrSerSer 1873
QY 1681 CCACCCCGCAGAGCTGCCAGAGTCCCTCAGACCTCCCGACGCTCCCGACAGATTCCTCC 1740
Db 1874 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSer 1893
QY 1741 AGTGGCTTCTCCACGAGCTGCGAGAGCTCAGACATGCTCATACGCGCGTCTGCC 1800
Db 1894 SerValProPheSerThrAspCysGlnAspSerAspMetLeuHisThrProGlyAlaAla 1913
QY 1801 TCCGACAGCTGACATGACTGACCCCATGAGAGCGGAGAGACCTTACGAGCAGCTG 1860
Db 1914 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeu 1933
QY 1861 CAGCCCGTGAATCTCCGCTCAGAGACCTCATTTGGGTCCCAAGGCGCGCTGAGAT 1920
Db 1934 GlnAlaValLysIleLeuAlaValArgAspLeuIleThrValProArgArgGlyGlyAsp 1953
QY 1921 GTTATGCTATGAGCTGGAGAGATTTCTGAAGCCACGCGGCGGAGATCATTTGGCGCTC 1980
Db 1954 ValIleValIleGlyLeuGlnLysAspSerGlyAlaGlnArgGlyArgValIleAlaVal 1973
QY 1981 TTAAGCCCGGAGAGCTGACTCCGATGGGAGTGTGTGATGCTGCCGTGGCAGAG 2040
Db 1974 LeuLysAlaArgGlnLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 1993
QY 2041 GACACTGTGTGTGACCTTTGAAATGAAGAAACACAGAGTGTGCTGCCCTGTGGAG 2100
Db 1994 AspThrValValCysThrPheGlnAsnGlnAsnThrGlnTrpCysLeuAlaValTrpArg 2013
QY 2101 GCGTGGGCGCGAGGAGTTCGACATTTCTACAGCTCTACAGAGAGCTGGCGCGCTG 2160
Db 2014 GlnTrpGlyAlaArgGlnPheAspIlePheTyrGlnSerTyrGlnGlnLeuGlyArgLeu 2033
QY 2161 GAGGCTTGCACTGCAAGAGA 2181
Db 2034 GlnAlaCysThrArgLysArg 2040
RESULT 6
US-10-143-133-2
; Sequence 2, Application US/10143133
; Publication No. US20020197658A1
; GENERAL INFORMATION:
; APPLICANT: yoganathan, Thillainathan
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use
; FILE REFERENCE: KINE-023
; CURRENT APPLICATION NUMBER: US/10/143,133
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,555
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0


```

Db 750 ProglythrProglythr-----ProArgSerPro 759
OY 1538 -----GCCACAGCGCCACCAAGAGTGCCTGAGGGGAGCTCCATCGCGAGC 1585
Db 760 ProleuglyleuileSerArgPro-----ArgProSerProleuArgSer 774
OY 1586 TGAGCATCATGATACAGAGAGTGGGACGACAGATCCTGATCCACAGATCACTCA 1645
Db 775 -----ArgIleasProTrp-----Ser 780
OY 1646 CTGACTACTGCTCCATGTCCTCTCTACTCTCATCCCAACCCCGCAGGCTGCCAGTCCC 1705
Db 781 PheValSerIle-----GlyProArgProSerProleuProSerPro 794
OY 1706 CCTCAAGCTCCCGACGTCGCCAGCAAGTCTCTCAGAGTGCCTTTCACACCGACTGG 1765
Db 795 -----GlnProIleProArgArgala---ProTrpThrLeuPheProSerSer 810
OY 1766 AGGACTCAGACATGCTACATACAGCCCGGTGCTCCGACAGTCTGTGAGCATGACTGA 1825
Db 810 -----810
OY 1826 CCCCATGAGCGGAGAGACT---TCAGCCAGACCTCGAGCGCTGAAGATCTCGCCG 1882
Db 811 ProPheTrpAspSerProProIleAsn-ProPheGlnGlyGlyProGlnAspCysArgAl 830
OY 1883 TCAGAGACCTCATTTGGGTCGCCAGCGCGGTGG 1916
Db 830 aglnThrIysAspMetGlyAlaGlnAlaProTrp 841

RESULT 7
US-09-771-161A-255
; Sequence 255, Application us/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 255
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-255

Alignment Scores:
Pred. No.: 1,91e-09 Length: 1036
Score: 291.00 Matches: 199
Percent Similarity: 33.57% Conservative: 88
Best Local Similarity: 23.27% Mismatches: 264
Query Match: 7.17% Indels: 304
Db: 10 Gaps: 44

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-771-161A-255 (1-1036)
OY 13 CTGCGGGGACACCGCATGAGAACTTCTCGAGTTCGCGAGAGCGCCAGCATGCTG 72
Db 165 LeuSerGlnProGlnAlaMet-----AspAspPheIleArgGlnValAsnAlaMet 181
OY 73 CAGCGCCTGAGACACCCCTGATCGTGGCGCTCATGGGATCGATC---CACCGGCTC 129
Db 182 HisSerLeuAspIleHisArgAsnIleArgLeuTrpGlyValValLeuThrProProMet 201
OY 130 TGCTTGCCCTGTGAGCTCGCGCGCTCAGCAGCATCAACACGCTGCTGTCGAGAACGC 189

```

```

Db 202 LysMetValThrGlnLeuAlaProLeuGlySerLeuLeuAspArgLeuArgLysHisGln 221
OY 190 AGAGATCTCTCTTATACCCCTGGAGACATGCTC-----ACCAGAAATAATAGCC 240
Db 222 -----GlyHisPheLeuLeuGlyThrLeuSerArgTrpAla 233
OY 241 TACAGATCGCTCGGCGCTGCTGCTACCTGCACAGAAACATCATCTTGTGACCTG 300
Db 224 ValGlnValAlaGlnGlyMetGlyTyrLeuGlnSerLysArgPheIleHisArgSerLeu 253
OY 301 AAGTCGACAAACATTTGCTGTGTGCTTGCCTTGACGTCAAGAGACACATCAATCAAGSTA 360
Db 254 AlaAlaArgAsnLeuLeuAlaThrArgAspLeu-----ValLysIle 268
OY 361 TCTGACTACGGGATTTGAGG-----CAGTCATTCCATGAGGGGCGCTTAGGC 408
Db 269 GlyAspPheGlyLeuMetArgAlaLeuProGlnAsnAspArgHisTyrValMetGlnGlu 288
OY 409 GTGAGGCGACTCCT---GGCTACACAGCCCGCAGAGATCAGGCTCGCATGTATATGAT 465
Db 289 HisArgLysValProPheAlaTrpCysAlaProGlnSerLeuLysThrArgThrPheSer 308
OY 466 GAGAGGTAGATATGTTCTCTTGAATGCTCTTACAGAGTCTGTCA---GGACAG 522
Db 309 HisAlaSerAspThrTrpMetPheGlyValThrLeuTrpGlnMetPheThrTyrGlyGln 328
OY 523 CGCCCTGCACGTGGCGCACACACAGCTCCAGATGCGCAAGAGCTGCTCAAG---GGCATC 579
Db 329 GluProTrpIleGlyLeuAsnGlySerGlnIleLeuHisLysIleSprGlyGluGlu 348
OY 580 CGCCCGGTTCTGGGCGACCCGAGAGAGTGCAGTTCCGCGCATCTCAGCGCTCATGATG 639
Db 349 Arg-----LeuProArgProIleAspCysPro---GlnAspIleTyrAsnValMetVal 365
OY 640 GAGTCTGGGACACTAAGACACAGAGACGACCGCTGCGCTCGGTGTCAGCCAGATG 699
Db 366 GlnCysTrpAlaHisLysProGlnAspArg-----375
OY 700 AAGGACCGACTTTGGCCACCTCATGATGACTGTGTGGAGAGCAGACGCTTC 759
Db 376 -----ProThrPheValAlaLeuArgAspPheLeuGlnAlaGlnProThrAspMet 393
OY 760 TTTCTATCCAGCGCAGAGAG-----TAC 783
Db 394 ArgAlaLeuGlnAspPheGlnGluProAspLysLeuHisIleGlnMetAsnAspValIle 413
OY 784 ACCGTGGTGTGGGATGGAAGAAGAGAGTCCAGAGACTACACGGGTGTAACACAGAG 843
Db 414 ThrValIle-----GlnGlyArgAlaGlnAsnTyrTrpTrpArgGlyGlnAsnThr 430
OY 844 AAGGCGCTCATGAGAGTGCAGAGATGCTGCTGCGCTGG-----ATGAAG 888
Db 431 -----ArgThrLeuCysValGlyProPheProArgAsnValVal 443
OY 889 GTGAGCTCCAGCTCCAGAGTCCAGAGATCCTGTGAGAGCCACCGCAGG-----937
Db 444 ThrSer-ValAlaIleGlyLeuSerIleArgGlnAspIleSerGlnProLeuGlnAsnSerPheI 463
OY 938 ---ACCAGAAATCTATCATCTACACCTCAAGGAGGATGCGCTTAAACACACCCCA 993
Db 463 eHisThrGlnHisGlyAspSerAspPro---ArgHisCysTrp-----GlyPheProAs 480
OY 994 CAGGCTTGATATCTCCAGCTGCTCAGCTGCTTGTGGCGCTGCTTATTAAGAAG 1053
Db 480 PArgIleAspGlnLeuTyrLeuGlnLysnProMetAspProProAspLeuSerValI 500
OY 1054 AATT-----CCTACCTGCTCTTA 1071
Db 500 uLeuSerThrSerArgProProGlnHisLeuGlyValLysLysProThrThrTyrTrpSpr 520
OY 1072 GCGGCGCTCG-----1081

```


QY 2071 AACACAGATGTCCTGCGCCCTCTGAGAGGCGCTGGGCGC 2111
 DB 857 --TyrTyrLeuLeuProGluArgProSerTyrLeuGluArg 869

RESULT 9
 US-10-014-882-2
 Sequence 2, Application US/10014882
 Patent No. US20020107384A1
 GENERAL INFORMATION:
 APPLICANT: Hu, Yi
 APPLICANT: Kieke, James
 APPLICANT: Donoho, Gregory
 TITLE OF INVENTION: NO. US0020107384A1 Human Kinase and Polynucleotides Encoding
 FILE REFERENCE: LEX-0279-USA
 CURRENT APPLICATION NUMBER: US/10/014,882
 PRIOR FILING DATE: 2001-12-11
 PRIOR APPLICATION NUMBER: US 60/254,744
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 1036
 TYPE: PRM
 ORGANISM: homo sapiens
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (1)...(1036)
 OTHER INFORMATION: Xaa - Any Amino Acid
 US-10-014-882-2

Alignment Scores:
 Pred. No.: 3,26e-09 Length: 1036
 Score: 287.00 Matches: 196
 Percent Similarity: 34.85% Conservat: 95
 Best Local Similarity: 23.47% Mismatches: 284
 Query Match: 7.07% Indels: 261
 Gaps: 36

US-09-836-392-8_COPY_22_2205 (1-2184) x US-10-014-882-2 (1-1036)

QY 25 GATGCCATGAGAACTTCGAGTTCCGGCAGAGGCGCAGATGTCAGCGCTGCAG 84
 DB 160 AspalalalalalalalaguserValArgArgGluAlaArgLeuPheAlaMetLeuArg 179
 QY 85 CACCCCTGATCGTGGCGCTCATCGGCATCAGCATCCACCGC-----CTGCTTCGCC 138
 DB 180 HisProAsnIleIleIleGluLeuArgGlyValCysLeuGlnGlnProHisLeuGlyLeuVal 199
 QY 139 CTGGAGCTGGCGCGCTCAGACAGCCCTCAACACCGCTGCTG---TCCGAGAACGCC----- 189
 DB 200 LeuGluPheAlaArgGlyGlyAlaLeuAsnArgAlaLeuAlaAlaAsnAlaAlaPro 219
 QY 190 -----AGAGATCTCTCTTATACCCCTGGACACATGTCACG 228
 DB 220 AsprArgAlaProGlyProArgArgAlaArgArgIlePro---ProHisValLeuVal 238
 QY 229 CAAAAAATAGCTACACAGATCGCCTCGCGCTGCTTACCTGCACAGAAA----- 279
 DB 239 ---AsnTrpAlaValGlnIleAlaArgGlyMetLeuTyrLeuHisGlnGluAlaPheVal 257
 QY 280 AACATCATCTCTGTCGAGTGAAGTGGACACATCTGTGTGGTCCCTTGACGTCACG 339
 DB 258 ProIleLeuHisArgPheLeuSerSerAsnIleLeuLeu-----LeuGluLysIle 275
 QY 340 GAGCAGATCAACATC-----AAGCTATGCTACAGGATTCGAGGACG 384
 DB 276 GlnHisAspArgPheLecysAsnLysThrLeuLysIleThrAspPheGlyLeuAlaArgGlu 295
 QY 385 TCATTCATGAGAGGCGCCTAGGCGCTGAGGAGCATCTGTCGTCACAGGCCACAGATC 444
 DB 296 TrpHisArgThrThrLysMetSerThrAlaGlyThrTyrAlaTrpPheAlaProGluVal 315
 QY 445 AGGCCCTGCATGTATATGATGAGAGATGATATGTTCTCTATGAGATGTCGCTCTAC 504

DB 316 IleLysSerSerLeuPheSerLysGlySerAspIleTrpSerTyrGlyValLeuLeuTrp 335
 QY 505 GAGTGTGTCAGGACAGGCGCCCTGCACGCGCACACAGCTCCAGATTCGCG----- 558
 DB 336 GlnLeuLeuThrGlyGlnValProTyrArgGlyIleAspGlyLeuAlaIleArgGly 355
 QY 559 -----AAGAGCTGTCAGAGGACATCCGCGCTTCGAGGAGCGAGGAACTG 609
 DB 356 ValAlaValAsnLysLeuThrLeuProIleProSerThrCysProGluPro----- 372
 QY 610 CAGTTCGCGGACTGCAGCGCTCATGATGAGTGTGTCGAGACACTAGCCAGAGAGCA 669
 DB 373 -----PheAlaLysLeuMetLysGlyCysTrpGlnGlnAspProHisIleArg 388
 QY 670 CCG-----CTGGCCCTG-----TCGGTGGAGACCG 696
 DB 389 ProSerPheAlaLeuLeuGlnGlnLeuThrAlaIleGlnGlyValAlaMetThrGln 408
 QY 697 ATGAAAGACCCGACTTTCGACCTTCATGATGAACTGCTGTCGGAAGCAGACAGCC 756
 DB 409 MetProGlnGlnSerPheHisSerMetGlnAspAspTrpLysLeuGluIleGlnMet 428
 QY 757 TTC-----TTCATCCGAGGCGCAGATACAC----- 786
 DB 429 PheAspGlnLeuArgThrLysGlyLysGlyLeuArgSerArgGlnGlnLeuThrArg 448
 QY 786 ----- 786
 DB 449 AlaAlaLeuGlnGlnLysSerGlnGlnGlnLeuLeuLysArgArgGlnGlnLeuAla 468
 QY 787 -----GTGGTGTTCGATGAGAA 807
 DB 469 GluArgGlnLeuAspValLeuGlnLysArgGlnLeuAsnIleLeuIlePheGlnLeuAsnGln 488
 QY 808 GAGGATCCAGGACACTACAGGCGGTGACACAGAGAGGCGCTCATGAGTGCAGAG 867
 DB 489 GlnLysProLys-----ValLysLysArgGlyLysPheLysArgSerArg 504
 QY 868 ATG---TGCTGCGCTGAGTGAAGTGAAGTGCAGCTGCAGCTCAGATCCAGAGA-----TCC 918
 DB 505 LeuLysLeuLysAspGlyHisArgIleSerLeuProSerAspPheIleHisLysIleThr 524
 QY 919 CTGTGACACACCGACGAGGACGAGAAATCTACATCTACACCTCAGGAGCATGTCGCC 978
 DB 525 ValGlnAlaSer---ProAsnLeuAspLysArgArgSerLeuAsnSerSerSerSer 544
 QY 979 TTAACACACCCAC-----AGCCTTGATACCTCCAGCTGCTCAGCTGCTTC 1029
 DB 544 OProSerSerProThrMetMetProArgLeuArgAlaIleGlnLeuThrSerAspLys 564
 QY 1030 T-----TGCCCGTGCCTGTATTA----- 1048
 DB 564 TrpAsnLysThrTrpGlyArgAsnThrValPheArgGlnGlnGluPheGlnAspValLysArg 584
 QY 1049 -----AAAAGATTCCTACTGCTGTAGCGGCGCTGCGGAGTGTGTGCT 1098
 DB 584 GAsnPheLysLysGlyCysThrTrp----- 593
 QY 1099 GTGTTCCCGTGTGCGGCGCACCCAAAGACAGCTGCTCTACTGTGCTCAGACACA 1158
 DB 593 ----- 593
 QY 1159 GCCAAGAGTCCAAATTCAGATCGGAGTGAAGACGACGAGAAACCTTACCAGTG 1218
 DB 594 -----GlyProAsnSerIle---GlnMetLysAspArgThrAspCysLysGlnArgGly 610
 QY 1219 AAGCCATGAGG-----TGCTCAACAGGCGCTGAGTGTGATACAC 1263
 DB 610 eArgProLeuSerAspGlyAsnSerProTrpSerThrIleLeuIleLysAsnGlnLysThr 630
 QY 1264 AATGGCGCGGCGCTCCTTG-----TCATGAGCTGTGCTCCTGAGATC 1308

```

Db 630 rmetProleuAlaSerLeuPheValAlaSpInProGlySerCysGluProLysLeuSe 650
QY 1309 TGCAGGCGGCTGGAGCCCTACATGCGCCCTCATGTTACTGATGCTGTCAGCTCT 1368
Db 650 rProAspGlyLeuGluHisArgLysPro----- 659
QY 1369 GAGGCGAGGAGGAGGAGGCTGCTGTCGCTGATGACAGAGCCACTCTGTTGATG 1428
Db 660 ----LysGlnIleLysLeuProSerGlnIleArgLysLeuProLeuGlyLysAspAl 678
QY 1429 TACCACTCCACCACTACCACTGTCGTGTCCTGATGACAGAGCCACTCTGTTGATG 1488
Db 678 aglnArgLysProAlaGlnAlaGlnSerTrpGlnGlnAlaAlaSerAlaAlaAla 698
QY 1489 AGGAGCATGTTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1548
Db 698 ArhValSerIleGluMetThrProThrAsnSerLeuSerAlaArgSerProGlnArgLysLys 718
QY 1549 AACCCAAAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1605
Db 718 sPhrGlnSerAlaLeuArgLys-----CysThrValLeuLe 730
QY 1606 -----GAGCTGGGAGCAGCAGATCTCTG---ATCCAC----- 1632
Db 730 uAlaSerValAlaLeuGlyLeuAspLeuArgLysLysLysAlaGlnAlaAlaGln 750
QY 1633 -----CAGGATCATCTCATCTGATGATGATGATGATGATGATGATGATGATG 1664
Db 750 uProLeuProLysGlnGlnLysLysLysLysLysLysLysLysLysLysLysLys 770
QY 1665 CTCCTACTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1706
Db 770 rArgArgSerAlaSerProProThrSerLeuProThrArgLysLysLysLysLys 790
QY 1707 CTCACAGCTCCCGAGCTCCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1760
Db 790 oProSerLeuProLeuSerSerAlaLeuGlyLysLeuSerThrProSerPheSerThr 810
QY 1761 CTGCGAG-----GA 1769
Db 810 scLysLeuGlnMetAlaSerSerLysLysLysLysLysLysLysLysLysLys 830
QY 1770 CTCAGACATCTACATACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1829
Db 830 pSerGlnMetLeu---ThrProAspPheCys---ProThrAla-----P 843
QY 1830 CATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1871
Db 843 rGlySerGlyArgGlnProAlaLeuMetProArgLysLysLysLysLysLysLys 863
QY 1872 GATCCCTGCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1931
Db 863 rGlnLeuProSerSerPheLeuGlnGlnThrCysGlyAsnValProTyrCysAlaSer 883
QY 1932 TGGCTGGAGA----- 1942
Db 883 erLysHisArgProSerHisHisArgArgThrMetSerAspGlyAsnProThrProThr 903
QY 1943 ----AGATTCTGAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1978
Db 903 LysAlaThrIleIleSerAlaThrGlyAlaSerAlaLeuPro 916

```

```

; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-19

Alignment Scores:
Pred. No.: 7,856-09 Length: 394
Score: 279.50 Matches: 87
Percent Similarity: 46.05% Conservative: 47
Best Local Similarity: 29.90% Mismatches: 108
Query Match: 6.89% Indels: 50
Gaps: 9

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-862-027-19 (1-394)
QY 7 AGGCACTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 66
Db 33 ArgHisAspProAspGlnAspLysSerGlnThrIleGlnValArgGlnGlnLys 52
QY 67 ATGCTGACAGGCGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 126
Db 53 LeuPheAlaMetLeuLysHisProAsnIleIleAlaLeuArgLysLysLysLysLys 72
QY 127 -----CTCTGCTTGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
Db 73 ProAsnLeuCysLeuValMetLupheAlaArgLysLysLysLysLysLysLysLys 92
QY 181 GAGAGCGCAGAGATCTCTTTATACCCCTGGGAGCAGATCTCACCAAAATA--- 237
Db 93 -----GlyLysArgIleProProAspIleLeu 101
QY 238 -----GCCTACAGATGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 279
Db 102 ValAsnTrpAlaValGlnIleAlaArgLysMetAsnTrpLeuHisAspGlnAlaIleVal 121
QY 280 AACATCATCTCTGTCGACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 330
Db 122 ProIleIleHisArgAspLeuLysSerSerAsnIleLeuIleLeuIleLysValGlnAsn 141
QY 331 ---GACGTCAAGAGGAGCAGATCAATCAAGTATGCTAGGAGGAGGAGGAGGAGGAGGAG 387
Db 142 GlyAspLeuSerAsnLysIle---LeuLysIleThrAspPheGlyLeuAlaArgGlnTrp 160
QY 388 TTCATGAGAGGCGCCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 447
Db 161 HisArgThrThrLysMetSerAlaAlaGlnIleThrAlaTrpMetAlaProGlnValIle 180
QY 448 CTTGCGATTGTATATGATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 507
Db 181 ArgAlaSerMetPheSerLysLysLysLysLysLysLysLysLysLysLysLysLys 200
QY 508 TTGCTGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 558
Db 201 LeuLeuThrGlyValValProPheArgGlyLysAspGlyLeuArgValAlaLysLysVal 220
QY 559 -----AAGAGCTGTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 612
Db 221 AlaMetAsnLysLeuAlaLeuProIleProSerThrCysProIlePro----- 236
QY 613 TTCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 672
Db 237 -----PheAlaLysLeuMetGlnAspCysTrpAsnProAspProHisSerArg--- 252
QY 673 CTGAGCCTGTGTGCTGTGAGCCAGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 732
Db 253 -----ProSerPheThrAsnIleLeuAspGln 261
QY 733 CTGTGCTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 791

```

```

Db 262 LeuThrThIleGluGluSerGlyPhePheGluMetProLysAspSerPheHisCysLeu 281
Qy 792 GTTTGGGATGGAAGAGAGCTCCAGAACTA 824
Db 282 GlnAspAsnTrpLysHisGluIleGlnGluMet 292

RESULT 11
US-09-862-027-43
/ Sequence 43, Application US/09862027
/ Patent No. US20020142428A1
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin R.
/ TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
/ FILE REFERENCE: 35800/234862
/ CURRENT APPLICATION NUMBER: US/09/862,027
/ PRIORITY FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 09/345,473
/ NUMBER OF SEQ ID NOS: 82
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO: 43
/ LENGTH: 626
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-862-027-43

Alignment Scores:
Pred. No.: 2,076-07 Length: 626
Score: 255.50 Matches: 150
Percent Similarity: 33.69% Conservative: 71
Best Local Similarity: 22.87% Mismatches: 242
Query Match: 6,308 Indels: 193
DB: 10 Gaps: 28

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-862-027-43 (1-626)
Qy 4 CTGAGGACCTCGGGGCCACCGATGCCATG-----AAGAAC 39
Db 74 VALATGHHASRGlnSerGlnValMetValLeuLysMetAsnLysLeuProSerAn 93
Qy 40 TTCTCCGAGTCCGCGAGGCGCAGCATGCGAGCGCTCGAGACCCCTGCATCGG 99
Db 94 ATGGGlyAsnThrLeuArgGlnValGlnLeuMetAsnArgLeuArgHisProAsnIleLeu 113
Qy 100 GCGCTATCGGATGAGCATCCACCGCGCTCTGCTCGCCCTGAGCTCGCGCTCAGC 159
Db 114 ArgPheMetGlyValCysValHisGln-----GlyGlnLeuHis 126
Qy 160 ACCCTCACACCGTGTCTGTCGAGAACGCCAGAGAT-----TCTTCTTTATACC 210
Db 127 ATAlaLeuThrGluTyrMetAsnGlyGlyThrLeuGlnGluLeuSerSerProGluPro 146
Qy 211 CTGGGACATGCTCAACCAAAATAGCTACAGATGCGCTCGGCGCTGGCTACG 270
Db 147 LeuSerTrpProValArgLeuHisLeuAlaLeuAspIleAlaArgGlyLeuArgTyrLeu 166
Qy 271 CACAGAAAAACATCATCTTCTGTGACCTGAGTGGAGACATTCCTGTGCTCCCT 330
Db 167 HisSerLysGlyValPheHisArgAspLeuThrSerLysAsnCysLeuValArgArgGln 186
Qy 331 GACGTCAGGAGACATCAACATCAAGCTATGTGACTACGGGATTTCCGAGCAG----- 384
Db 187 Asp-----ArgGlyPheThrAlaValAlaGlyAspPheGlyLeuAlaGluLysIlePro 204
Qy 385 TCATTCATGAGGCGGCC-----CTAGGCGTGAGGCGCATCTGGCTACGAG 432
Db 205 VALTyrArgGlnGlyAlaArgLysGluProLeuAlaValAlaGlySerProTyrTrpMet 224
Qy 433 GCCCAGAGATCAGGCGCTCGATGTATATGATGAGAAAGTAGATATTTCTCTATGA 492
Db 225 AlaProGluValLeuArgGlyLysLeuTyrArgGlyLysAlaAspValPheAlaPheGly 244
Qy 493 ATGGTGTCTTACGAGTGTGTCTCA-----GGACAGCGCCCTGCATG 534

```

```

Db 245 ILeuValLeuCysGluLeuLeuAlaArgValProAlaAspProAspTyrLeuProArgThr 264
Qy 535 GCCACACACCCACCTCCAGTTCCCAAGAGCTGTCCAGGCGATCCGCCGTTCTGGG 594
Db 265 GlnAspPheGlyLeuAspVal-----ProAlaPheArgThrLeuValGly 279
Qy 595 CAGCCGAGAGAGTGCAGTCCGCGGAGTGCAGGCGCTCATGATGAGTGGAGACT 654
Db 280 AspAspCysProLeuProHe-----LeuLeuAlaIleHisCysLysLeu 296
Qy 655 AAGCCAGAGACCGACCGCTGCGCTGTG-----GTGGTGACC 693
Db 297 GluProSerThrArgAlaProPheThrGluIleThrGlnHisLeuGluTyrLeuGln 316
Qy 694 CAGATGAGGAGCCGACATTTGCCACCTTCATGTATGAACTGCTGTGCGAGAGACA 753
Db 317 GlnLeuProGluProAlaProLeuThr-----ArgThr 327
Qy 754 GCGTTCTTCTCATCCAGGCGCAGAGTACACCGTGTGTTTGGAGTGAAGAGAG 813
Db 328 AlaLeuThrHisAsnGlnGly-----SerValAlaArgGly----- 339
Qy 814 TCCAGAGACTACACGCTGTGTGAACAGAGAGGCGCTCATGAGAGTGCAGATGTGC 873
Db 340 -----GlyPro----- 341
Qy 874 TGCCTGGATGAGGTAGTGCAGCTCCAGCTCCAGTCCAGAGATCCCTGTGGACACC 933
Db 342 -----SerAlaThrLeuProArgProAspProArgLeu-----SerA 354
Qy 934 GAGACAGAGAAATCTACATCTACACCTCA----- 965
Db 354 rGSerArgSerAspLeuPheLeuProProSerProGluSerProAsnTrpGlyAspA 374
Qy 966 -----GGCATGTGCCCC 978
Db 374 snLeuThrArgValAsnProPheSerLeuArgGlnAspLeuArgGlyLysIleLysL 394
Qy 979 TTAACACACCCACAGGCGCTGTGATCTCCAGTGTGCTGCTGCTGTGCGCTG 1038
Db 394 euLeuAspThrProSer--LysProAlaLeuProLeuValProProSerProPhePro 413
Qy 1039 CCTGTATTAAAGAAATCTTACCTGTGTAGCGGCGCTGCGAGTGGCGTGTGCT 1098
Db 413 rThrGlnLeuProLeuValThrThrProGluThrLeuValGlnPro----- 428
Qy 1099 GTGTTCCCGTGTGCGGCGCACCCCAAG-----ACAGTGTCTCTAC 1143
Db 429 -----GlyThrProAlaArgArgCysArgSerLeuProSerSerProG1 443
Qy 1144 CTGTGCTACACACAGCCACA-----GTGCCAG 1173
Db 443 uLeuProArgArgMetGluThrAlaLeuProGluProGluProAlaValAlaGlyPro 463
Qy 1174 TTCACATCGCGGATGAGAGCAGCGAGAGCCCTTACCCAGTGAAGGCGATGAGGTG 1233
Db 463 rAlaGlnGluMetGlnCysGlnGlySerSerProGluProGluProGluProAl 483
Qy 1234 GTCAACAGCGGCTGTGAGTGTGTACAGCAATGGCGGCGCTCTGTCTACAGTGT 1293
Db 483 aProGlnLeuProLeuAlaValAlaThrAspAsnPheIleSerThrCysSerSerLase 503
Qy 1294 GCGTCCCTG----- 1303
Db 503 rGlnProTrpSerProArgSerGlyProValLeuAsnAsnAsnProProAlaValAla 523
Qy 1304 AGATCTGAGGCGCGCTGAGGCGCTTACATGAGGCGCGCTCATGCTGAGTGTGCTGC 1362
Db 523 LAsnSerProGlnGlyTyrAla-----GlyLeuProTrpAlaArgAlaGlnHis 540
Qy 1363 ACCTGTGAGGCGAGAGGAGAGAGTGTGTGCTGCTGATGACAGAGGCGGCGCTGCT 1422

```

```

Db 540 rleuProArGala-----AlaAlaLeuGluArgThrGluPro-- 552
QY 1423 GGTATGTACACATCCACACATGCTGTGTGCGGCTACTTTCGGGCTCCAGC 1482
Db 553 -----SerProProSerAlaProAlaGluProAspGluGlyLeuProCy 568
QY 1483 CCCCTGAGGACATGTTCCGCTGCGGCTTGAGCGGAAACCCCGGACGACCAAC 1542
Db 568 sPro--GlyCysCysLeu-----GlyProPheSerPheGlyPheLeuSerMetCysPr 585
QY 1543 AGCGCCACACCAAGGTGCTGAGGGGAGCTCCATCGCGAGCTGAGCATCATGACAG 1602
Db 585 oArgProThrProAlaValAlaAlaArgTyrArgAsn-----LeuAsnCysGluAl 601
QY 1603 GAGGAGCTGGGACGAGATCCTGATCCACACGAAATCACTCACTGACTGCTCCATG 1662
Db 601 agGlySerLeuLeu-----CysHisArgGlyHisHisAlaLys----- 613
QY 1663 TCCTCTACTCTCTATCCACCCGCGGCGGAGCTGCGAGGT 1702
Db 614 -ProProThrPro-----SerLeuGlnLeuProGly 623

RESULT 12
US-09-836-081-2
; Sequence 2, Application US/09835081
; Patent No. US20020151020A1
; GENERAL INFORMATION:
; APPLICANT: VAN, Xianghe et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001224
; CURRENT APPLICATION NUMBER: US/09/835,081
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 752
; TYPE: PRT
; ORGANISM: Human
; US-09-835-081-2

Alignment Scores:
Pred. No.: 2, 12e-07 Length: 752
Score: 255.50 Matches: 183
Percent Similarity: 35.498 Conservative: 80
Best Local Similarity: 24.708 Mismatches: 229
Query Match: 6.308 Indels: 249
DB: 10 Gaps: 38

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-835-081-2 (1-752)
QY 13 CTGGGGGCGACCGATGCGATGAGAACTTCTCGAGTTCGGGCGAGGAGCGCAGCATGCTG 72
Db 95 LeuAsnProSerSerLeuGlnLysLeuPheArgGluValArg-----IleMet 110
QY 73 CACCGCTGCACACACCCCTGCATGTCGCGCTCATCGGCATC-----AGCATCCACCCG 126
Db 111 LysGlyLeuAsnHisProAsnHisLeuAlaLysLeuPheGlnLysLeuIleGluThrGluLysThr 130
QY 127 CTCGCTTCGCGCGCGAGCTGCGCGCTGACGACGCTCAACACCGCTGCTCCGAGAAC 186
Db 131 LeuTyrLeuValMetGluTyrAlaSerAlaGlyGluValAlaPheAspTyrLeuValSerHis 150
QY 187 GCCAGAGATTCCTCTTATACCCCTGGGACACATGCTCAACCAAAATAATGACC----- 240
Db 151 GlyArg-----MetLysGluLysGluAlaAlaArgAla 160
QY 241 -----TACCAATGCGCTCGCGCTGCGCTTACCTGCACCAAGAAAACATATCTTCTGT 294
Db 161 LysPheArgGlnIleValSerAlaValHisTyrCysHisGlnLysAsnIleValHisArg 180
QY 295 GACCTGAAGTGGGACACATTCGTGTGTGTCCCTTGACGTCAAGGAGACATCAATCATC 354

```

```

Db 181 AspLeuLysAlaGlnAsnLeu-----LeuAspAlaGlu-----AlaAsnIle 195
QY 355 AAGCATCTGATACAGGAGATTCGAGGACGATCATTCATGAGGCGCCCTAGGCTGGAG 414
Db 196 LysIleAlaAspPheGlyPheSer-----AsnGlnPheThrLeuGlySerLys 211
QY 415 -----GGCAGCTCTGGCTACACAGGCCCCAGAGATCAAGGCTCCGATTTGA 459
Db 212 LeuAspThrPheCysGlySerProProTyrAlaAlaProGluLeuPheGlnGlyLys 231
QY 460 TATGAT-----GAGAGGTAAATATGTTCTCTATGAAATGCTGTCTACAGTGTCTGA 516
Db 232 TyrAspGlyProGluValAspIleThrPheSerLeuGlyValIleLeuTyrThrLeuValSer 251
QY 517 GGACAGCGCCCTGACTGGAGCCACACAGCTCCAGATTCGCAAGAGGTCCAGAGCGC 576
Db 252 GlySerLeuProPheAspGlyHisAsnLeuLysGluLeuArgGluValLeuArgGly 271
QY 577 ATCCGCCCGGTT-----CTGGGCGACGCGGAGAGATGCAATTCGCGCA----- 621
Db 272 LysTyrArgValProPheTyrMetSerThrAspCysGlySerIleLeuArgArgPheLeu 291
QY 622 -----CTCCAGCGCCCTATGATGAGTGCCTGG----- 648
Db 292 ValLeuAsnProAlaLysArgCysThrLeuGlnGlnIleMetLysAspLysTrpIleAsn 311
QY 649 -----GACACTAAGCCAGAGACGACGCG----- 672
Db 312 IleGlyTyrGlnGluGlyLeuLysProTyrThrGluProGluAspPheGlyAsp 331
QY 673 -----CTGGCCCTGTCTG 664
Db 332 ThrLysArgIleGluValMetValGlyMetGlyTyrThrArgGluGlnLysGluSer 351
QY 685 GTGTGAGCCAGATGAGAGACCGGACTTTGCCACCTTCATGATGAACTGTCTGTGGG 744
Db 352 LeuThrSerGlnLysTyrAsnGluValThrAlaThrTyr-----LeuLeuGly 368
QY 745 AACGACAGACCTTCCTTCATCCGAGGCGGAGAGACACCGGTGTGGGATGA 804
Db 369 ArgLysThrGluGluGlyLysAspArgGlyAla-----ProGly 381
QY 805 AAGAGAGATCCAGAACTACACGCTGTGTAACACAGAGAGGCGCTCATGAGTGCAG 864
Db 382 LeuAlaLeuAlaLysValAlaGalaProSerAspThrThrAsnGlyThrSerSer-SerLys 401
QY 865 AGGATGTGT--GCCCTGGATGAGTGAAGTGCAGCTCC-----AG 906
Db 401 sGlyThrSerHisSerLysGlyGlnArgSerSerSerSerThrTyrHisArgGlnArg 421
QY 907 GTCCAGAGATCCCTGTGACAGACCCGAGGAGACCAAGAAATCTCATCTCAACCTCAAG 966
Db 421 GHisSerAspPheCysGlyProSerPro-AlaPro-----LeuHisPro 435
QY 967 GGCATGTGCCCTTAACACACACCCACAGAGCGCTTGATACCTCAGCTGTGCACCTGC 1026
Db 436 -----LysArgSerProThrSerThrGlyGlnAlaGlnLeuLys----- 448
QY 1027 TTTCTGGCCGTGCTCTTATTAAGAAATTCATACCTGCTTACGGGCGCTCCGAT 1086
Db 449 -----GluGlnArgLeuProGly-----ArgLysA 457
QY 1087 GGGCTTGTGCTGTGT--TCCGCTGTGCGGGGACCCCAAGAGACAGTCTCTCTAC 1143
Db 457 LAspCysSerThrAlaGlySerGlySer-ArgGlyLeuProProSerSerProMetVal 476
QY 1144 CTGTGCTACACACAGCAGCAGGTCCCAAGTTCAGATCGCGGATGAAGACGACGACGAG 1203
Db 477 SerSerAlaHisAsnProAsnLysAlaGlnIleProGluArgGlyLysAspSerThrSer 496
QY 1204 AACCCCTAACCCAGTGAAGGCGCATGAGGTGTCAACAGCGGCTGTGAGTCTGTACAGC 1263

```



```

Db 335 -----SerValProArgLgLygProSerAlaThrLeuProArgSerAProArgL 352
QY 895 TGCCA-----GCTCCAGTCCAGAGATCCCTGGAGC----- 926
Db 352 euserArgSerArgSerApleuPheleuProProSerProGluSerProProSerTrpG 372
QY 927 -----AGCCACCAGGAGCCAGAAATCTACATCTACACCCCTCAAGGC----- 969
Db 372 LyAspAsnLeuThrArgValAsnProPheSerLeuArg-GluAspLeuArgGlyLyLys 391
QY 970 -----ATGCGCCCTTAAACACACCCCAA 993
Db 392 IleuLeuLeuAspThrProCysLysProAlaThrProLeuValAlaProProSer 411
QY 994 CAGCGCTTGATACACCGAGCTGCTCACTGCTGCGCGCTGGCTGCTGTTTAAAG 1053
Db 412 ProLeuThrSerThr-GlnLeuProLeuValAlaSerProGluSerLeuValGlnProG 431
QY 1054 AATTCCTACCTGCTGTAGCGGCGCTCGCGATGGGCTTGCTGTGCTTCCGCTGCTG 1113
Db 431 UThrProValArgArgCysArgSerLeuProSerSerProGluLeu----- 446
QY 1114 CGGGCACCACCAAG-----ACAGCTGCTCTACCTGTGCTCAGACACCAAGCAG 1167
Db 447 -----ProArgArgMetGluThrAlaLeuProGly---ProGlyProSerProValG 463
QY 1168 TCCAGTTCAGATCGCGGATGAAGACGAGCAGAACCCCTACCACTGAAGGCATG 1227
Db 463 YProSerThrGluGluArgMetAspCysGluGlySerSerProGluProGluProProG 483
QY 1228 GAGTGTGTCACACGCGGCTGTGAGTGTGTGTACAGCAATGGCGCGCTCTGTTCATC 1287
Db 483 YProAlaProGluLeuProLeuAlaValAlaThrAspAsnPhelIleSerThrCysSer 503
QY 1288 GACTGTGCTCCCTGAGATCTGACAGCGCGCTGAGCCCTACATGCCCCCTCCA----- 1342
Db 503 rAlaSerGlnProTrpSerAlaArgProGlyProSerLeuAsnAsnProProAlaVal 523
QY 1343 -----TGGTACGTACGCTGTGACAGTCTGACAGGCTGTGAGGAGAGAGG 1380
Db 523 lValValAsnSerProGlnGlyTrpAlaArgGluProTrpAsnArgAlaGlnHisSer 543
QY 1381 GAGGAGTGTGTGTGCTGTGATGACAGGCCAACCTCTGTGTGTATGACACTCCACC 1440
Db 543 UProArgAlaAlaAlaLeuGlnUArgThrGluProSer-----Pr 556
QY 1441 ACCTACCACTGTGTGCGGCTACTTCTCGGGGTCCCAAGCCCTCAAGGACATGTT 1500
Db 556 oProProSerAlaProArgGluGlnGluGlnLeuProCysPro--GlyCysCysLe 575
QY 1501 CCGGTGCGGCTTGAGACGGAACCGCGGAGCCAGCCACACAGGCAACCAAGGTG 1560
Db 575 U-----SerProPheSerPheGlyPheLeuSerMetCysProArgProThrProAlaVal 593
QY 1561 CCTGAGGGGAGCTCCATCCGCGAGCTGACATCATGATGAGAGAGAGGCGAGCAG 1620
Db 593 lAlaArgTrpArgAsn-----LeuAsnCysGluAlaGlySerLeuLeu----- 607
QY 1621 AATTCGATCCACCAAGATCACTCACTGACTGCTCATGCTCTCTCTCTCTCTCTCATCC 1680
Db 608 -----CysHisArgGlyHisHisAlaLys-----ProProThrPro----- 619
QY 1681 CCACCCCGGAGCTGCGAGGT 1702
Db 620 -----SerLeuGlnLeuProGly 625

```

```

FILE REFERENCE: PA-0037 P
CURRENT APPLICATION NUMBER: US/09/974,298
CURRENT FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,331
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO: 129
LENGTH: 937
TYPE: PR
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID NO. US20020156263A1 1331526CD1
US-09-974-298-129

Alignment Scores:
Pred. No.: 1,02e-06 Length: 937
Score: 244.00 Matches: 100
Percent Similarity: 37.56% Conservative: 66
Best Local Similarity: 22.62% Mismatches: 157
Query Match: 6.01% Indels: 119
DB: 9 Gaps: 15

US-09-836-392-8_COPY.22_2205 (1-2184) x US-09-974-298-129 (1-937)
QY 4 CTGAGGCACTGCGGCGCCAGCCAGTGCATGAAGAACTTCCGAGTCCGCGAGGAGCC 63
Db 505 IleuYsThrLeuLeuAspTrpAsnAsnProGlnGlnTrpMetGlnPheGlnGlnAla 524
QY 64 AGCATGTGACAGCGCTCGACGACCCCTGCATCTGTGCGCTCATGCGC-----ATCAGC 117
Db 525 SerLeuMetAlaGlnLeuHisHisProAsnIleValCysLeuLeuGlyAlaValThrGln 544
QY 118 ATCCACCGGCTGTGC-----TTCCGC 138
Db 545 GluGlnProValCysMetLeuPheGluTrpIleAsnGlnGlyAspLeuHisGluPheLeu 564
QY 139 CTGAGCTGCGCGCCCTGACAGCAGCTCAACCGCTGTGCTGCGAGACCGCAGATTCCT 198
Db 565 IleMetArgSerProHisSerAspValGlyCysSerSerAspGluAspGlyThrValLys 584
QY 199 TCCCTTATACCCCTGGAGACATGCTCAACCAAAATAGCTTACCAAGTCCCTCGGCGC 258
Db 585 SerSerLeuAspHisGlyAspPheLeu-----HisIleAlaIleGlnIleAlaIleGly 602
QY 259 CTGGCTTACCTGCACAGAAGAAACATCACTCTGTGAGCTGAGCGAGACAACTTCG 318
Db 603 MetGluTrpLeuSerSerHisPhePheValHisLysAspLeuAlaAlaArgAsnIleLeu 622
QY 319 GTGTGTCCTTGAGCTCAAGAGACATCAACATCAACCTATCTGCTACCTACGAGATTCG 378
Db 623 -----IleGlyGlnGlnLeuHisLysAlaLysIleSerAspLeuGlyLeuSer 637
QY 379 AGCGATCATTCATGAGGCGCCCTAGCGGTGGAGGCACTCT-----GCC 426
Db 638 ArgGluIleTyrSerAlaAspTrpTrpArgValGlnSerLysSerLeuLeuProIleArg 657
QY 427 TACCAAGCCCAAGATCAAGGCTGCATGTATATGATGAGAGATAGATATGTTCTCC 486
Db 658 TrpMetProProGluAlaIleMetGlyLysPheSerSerAspSerAspIleTrpSer 677
QY 487 TATGAAATGCTCTACAGATGTGTCTCA-----GACACCGCCCTGACATGGGCAACAC 543
Db 678 PheGlyValValLeuTrpLulPheSerPheGlyLeuGlnProTrpTrpGlyPheSer 697
QY 544 CAGCTCCAGATTGCAAGAAGCTTCCAAAGGCAATCCGCCGCTTGTGGGACCGCGAG 603
Db 698 AsnGlnGluValIleGluMetValArgLys-----ArgGlnLeuLeuProCysSerGlu 715
QY 604 GAAGTGCATTCGCGGAGATCGAGCGCCATCATGATGAGATGAGTGGGAGCACTAAGCCAGAG 663
Db 716 AspCysPro--ProArgMetTyrSerLeuMetThrGluCysTrpAsnGluIleProSer 734

```

```

RESULT 14
US-09-974-298-129
Sequence 129, Application US/09974298
Patent No. US20020156263A1
GENERAL INFORMATION:
APPLICANT: Chen, Hwei-Mei
TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER

```

```

QY 664 AACGACCCGCTGGCCTGTCGGTGTGAGCCAGATGAGACCCGACTTTGCCACCTTC 723
DB 735 ArgArgProArgPheLysAspIleHisValArgLeuArgSer----- 748
QY 724 ATGTATGAACTGTGTGTGGAGACAGACAGCTTCTTTCATCCAGGCGCAGAGTAC 783
DB 748 ----- 748
QY 784 ACCGTGTGTTTGGATGGA----- 748
DB 749 -----TrpIuGlyLeuSerSerHisThrSerSerThrProSerGlyGly 764
QY 820 AACTACACCGGTGTGAACACAGAGAGGCGCTATGAGGTGCAGAGATGCTGCTCCCT 879
DB 765 AsnAlaThrThrGlnThrThrSerLeuSerAlaSerProValSerAsnLeuSerAsnPro 784
QY 880 GGGATGAGGTGAGTGCAGCTCCAGTCCAGATCCCTGTGAGACAGCCAGAGAC 939
DB 785 -----Arg 785
QY 940 CAGAAATATCATCTACATCAACCTCAAGGCGATGCTTAAACACACCCCAAGAGCC 999
DB 786 TyrProAsnTyrMetPheProSerGlnGlyLe-----ThrProGln----- 799
QY 1000 TTGGATACCCAGCTGCTCTACCTGCTTCTTGCGCGCTGTTATTAAGAATTCC 1059
DB 800 -----GlyGlnIleAlaGlyPheIleGlyProProIleProGlnAsnGlnArg 815
QY 1060 TACCTGTGTAGACGGCGCTCCGCGATGGGCTGTGTGCTGTGTTCCGCGTGGCGGGC 1119
DB 816 PheIleProIleAsnGlyLys-ProIle----- 824
QY 1120 ACCCAAGACAGACGCTCTTACCTGCTGCTACACAGACCAAGAGT----- 1168
DB 825 -ProProGlyTyrAlaAlaPheProAlaAlaHisTyrGlnProThrGlyProProArgVal 844
QY 1169 -----CCAAGTCAAGCATCGCGGATGAGAGACCGCAGAGACCC 1209
DB 844 ILeGlnHisCysProProProLysSerArgSerProSerSerAlaSerGlySerThrIse 864
QY 1210 TACC 1213
DB 864 rThr 865

RESULT 15
US-09-977-260-6
/ Sequence 6, Application US/09977260
/ Publication No. US20020192790A1
/ GENERAL INFORMATION:
/ APPLICANT: ULLRICH, AXEL
/ APPLICANT: GISHIZKY, MIKHAIL
/ APPLICANT: SURES, IRMINGARD
/ TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
/ FILE REFERENCE: 038602/1260
/ CURRENT APPLICATION NUMBER: US/09/977,260
/ PRIOR APPLICATION NUMBER: 08/232,545
/ PRIOR FILING DATE: 1994-04-22
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 505
/ TYPE: PRT
/ ORGANISM: Unknown Organism
/ FEATURE:
/ OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
US-09-977-260-6

```

Alignment Scores: 1.01e-06 Length: 505
 Pred. No.: 243.50 Matches: 75

```

Percent Similarity: 50.41%
Best Local Similarity: 30.49%
Query Match: 6.00%
DB: 9
Gaps: 11
Conservative: 49
Mismatch: 77
Indels: 45

US-09-836-392-8_COPY_22_2205 (1-2184) x US-09-977-260-6 (1-505)

QY 43 TCCGAGTTCGGCAGAGAGGCGCAGCATGTGACAGCGCTGCAGACCCCTGCATCTGGCG 102
DB 272 AsnAspPheLeuArgIleValGlnIleMetLysAsnLeuArgHisProLysLeuIleGln 291
QY 103 CTCATCGGATCAGCATC-----CACCGCTCTGCTTGGCCCTGGAGCTCGCGCGCTC 156
DB 292 LeuTyrAlaValCysThrLeuGlnAspProIleTyrIleThrGlnLeuMetArgHis 311
QY 157 ACAGCCTCAACACCGCTGCTCCGAGAACGCCAGAGATTCTCTTATACCCCTGGGA 216
DB 312 GlySerLeuGlnGlyLysLeu-----GlnAsnAspThrGlySerLysLe----- 326
QY 217 CACATGCTACCCCAAAATA-----GCCATCAGATGCGCTGGCGCTGCTACCTG 270
DB 327 ---HisLeuThrGlnGlnIleValAspMetAlaAlaGlnIleValAlaSerGlyMetAlaTyrLeu 345
QY 271 CACAGAAATAATCATCTCTCTGTGTGACCTGACCTGAGTGGAGACAATCTGCTGTGCTT 330
DB 346 GluSerArgAsnTyrIleHisArgAspLeuAlaAlaArgAspValLeu----- 361
QY 331 GACCTCAAGACACATCATCAACATCATGATCTGATGATGAGATTTCGAG----- 381
DB 362 ---ValGlyGlnHisAsnIleTyrLysValAlaAspPheGlyLeuAlaArgValPheLys 380
QY 382 -----CAGTCATTTCATGAGGCGCGCTGAGCGCTGAGGCGACT 420
DB 381 ValAspAsnGlnAspIleTyrGlnSerArgHisGlnIleLysLeuProValLysTrpThr 400
QY 421 CCGGTGTACAGAGCCCGCAGAGATCAAGCGCTCCCATGTTATGATGAGAGATATAG 480
DB 401 -----AlaProIleAlaIleArgSerAsnLysPheSerIleLysSerAspVal 416
QY 481 TTCCTGTATGATGTGCTTACGAGTTGCTGTCA---GGACAGGCGCTGCTGAGCGGC 537
DB 417 TrpSerPheGlyIleLeuLeuTyrGlnIleIleThrTyrGlyLysMetProTyrSerGly 436
QY 538 CACACACAGCTCCAGATTGCCAAGAGCTGTCCAAAGGCGATCCGCGGCTGTGGGCGAG 597
DB 437 MetThrGlyAlaGlnValIleGlnMetLeuAlaGlnAsnTyrArg-----LeuProGln 454
QY 598 CCGGAGGAGTGCAGTTCGGGCGACTGCAGGCGCTCATATGATGATGCTGGAGACTAG 657
DB 455 ProSerAsnCysPro---GlnGlnPheTyrAsnIleMetLeuGlnCysTrpAsnAlaGln 473
QY 658 CCAGAGAACGACAGCGCTGCGCTGTGCGGTGTGAGCCAGATGAGAGACCCGACTTTTCC 717
DB 474 ProLysGlnArg-----ProThrPheGln 481
QY 718 ACCCTCATGTATGAAGT 735
DB 482 ThrLeuArgTrpLysLeu 487

```

Search completed: April 15, 2003, 07:24:33
 Job time : 91 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 15, 2003, 07:21:48 ; Search time 3444 Seconds

(without alignments)
6151.813 Million cell updates/sec

Title: US-09-836-392-21

Sequence: 1 MURHLRATDAMKNFSEFROE.....IFYQSYELGRLEACTRRKR 728

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=rlh
-Q/cg2_3/USPTO.spool/US09836392/runat_08042003_090540_23400/app-query.fasta_1.903
-DB=GenEmbl -OFFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIDE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09836392.ccdn_1.1.2189.ccdn_08042003_090540_23400 -NCPU=6 -ICPU=3
-NO_ALPHY -NO_MAMP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_in:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pin:*
35: em_hgt_rtd:*
36: em_hgt_mam:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgtgo_hum:*
40: em_hgtgo_mus:*
41: em_hgtgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3822	100.0	5370	9	AB058693
2	3811	99.7	4210	9	AK058028
3	3579.5	93.7	6045	6	AX250161
4	3578.5	93.6	3509	6	BC005408
5	3225	84.4	3112	9	AK054662
6	1961	51.8	2790	9	AK094556
7	1653.5	43.3	1980	10	BC027199
8	1567.5	41.0	110000	2	AL390202_01
9	1567.5	40.8	178529	2	AL390202_02
10	1560.5	40.8	178876	2	AC090907
11	1560.5	40.8	178876	2	AC019254
12	1560.5	40.8	183401	2	AC020578
13	1352	35.4	2047	9	AK026772
14	1295	33.9	288428	2	AC127595
15	856	22.4	526	6	AR175612
16	582.5	15.2	179377	2	AC117156
17	479.5	12.5	80468	2	AC096233
18	404	10.6	97358	2	AC017943
19	404	10.6	170939	3	AC007757
20	404	10.6	197071	3	AC007771
21	404	10.6	225804	3	AE003731
22	376	9.8	32556	3	AF098504
23	330.5	8.6	22730	6	AX165563
24	326.5	8.5	3030	10	BC034074
25	315	8.2	1631	8	AY085535
26	315	8.2	1795	8	AY125513
27	314	8.2	1267	8	AY133876
28	314	8.2	1798	8	AY046026
29	311	8.1	3138	9	HSMTMR
30	307.5	8.0	3454	6	AX337846
31	307.5	8.0	3454	9	HARNAMK2
32	307.5	8.0	179377	2	AC117156
33	299	7.8	3616	10	BC021891
34	292	7.6	6436	3	HSMTMR
35	288.5	7.5	3447	3	AF416233
36	288.5	7.5	5129	3	AY045717
37	287	7.5	1805	8	AY036609
38	285.5	7.5	38340	2	F6E21
39	285.5	7.5	194920	8	ATCRIRV74
40	281.5	7.4	167054	3	AP004744
41	279.5	7.3	5131	3	AY119549
42	278.5	7.3	1672	8	AY096470
43	278.5	7.3	1672	8	AY070086
44	278	7.3	3931	9	AF251442
45	275.5	7.2	8125	3	AF481923

RESULT 1

ALIGNMENTS

AB058693
 LOCUS AB058693 5370 bp mRNA linear PRI 05-JUN-2001
 DEFINITION Homo sapiens mRNA for KIAA1790 protein, partial cds.
 ACCESSION AB058693
 VERSION AB058693.1 GI:14017796
 KEYWORDS
 ORGANISM Homo sapiens brain cDNA to mRNA, clone:fh24104.
 Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (sites)
 Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R. and Ohara, O.
 Prediction of the coding sequences of unidentified human genes. XX.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro
 DNA Res. 8 (2), 85-95 (2001)
 21245130
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute,
 Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,
 Fax: 81-438-52-3914)
 FEATURES
 source
 1..5370
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fh24104"
 /rissue_type="brain"
 /note="vector: pBluescriptII SK plus"
 1..5370
 /gene="KIAA1790"
 <1..4112
 /gene="KIAA1790"
 /note="Start codon is not identified."
 /product="KIAA1790 protein"
 /protein_id="BAB47419.1"
 /db_xref="GI:14017797"
 /translation="PROGSTLELELQGRAPQVHGEATITTWELORPAGSRKAV
 ESSEFNVMDIGPAMATVNOCFETDKALYYVNMALAEENAVANQEWLNIKAP
 NAVVYVGLHDLIAKFEVERIALRLAYVLAICSPSGSRATGPDITFKLHISC
 KSLGEOGLROLIFHTCSMDVSGTIGORLAGRLIPRSYSLDEAVLAEOGRSD
 DDVOYLTRQLEOLYEOTPDNDIKYEDQSAISPLIENGTLHPDHSGRNLYEL
 DPIMISELQRIEFNFKSGSVAKNGVIRAEEDRLMLVGTGFTQCTEBOYFOLAFEI
 ALPVANDSTLPLHLPSKRGDTHGMRHETANTIQRFEMSVVPVGFWORLARNLIS
 LAEMOLOLEFENKKNKTSRNKRYTIIISFGNQRNRSTEFKMSVPIYMOEGLLVTEFG
 GYLSESSDVNMKKKSGMKIVCOSEVDYFAMAFITDHNLSLIDOMPALATIESD
 GTPLEOYVPCVPCETAMAOHTDPSEKSEDDVOYFMEHDCVLAIRBDFISCRHDP
 VPIOLVPELPEMTDPAFLFLENSKLESEDESGVYAGOGSCTVYRAAOCPYAVK
 RPHIKFKNFANVPADPLMRHLRATDAKNSEFROEASMLALGHPCTVALIGLSIH
 PLCELELAPLSLNTVLSENARSSFLPLGMLTQKLAQIASGLAHLKNNILIFCD
 LKSDNHLVSLDVKHEHINKLSDYGISRQSFHEGALVGEYGTQGLAELRPVYADEK
 VDMFSGVWLYELLEGORPALGHOLAKLSKISIRPVLAGPEEVOYFRLDALMEK
 WDTEPERPLIASVYOSKMDPFEATEMEVLCGKOTAFSSGOEYVTFVFMGKESR
 NYTVNTEKGLMEVOMCCPGMKVSCOLOVOBSIMTAEEDOKIYITLTKGMPLMPO
 CALDTPAVYTCFLAVPYIKNSYLVLAGLADLVAVFPYRRTPTPDDSCYLCSTRANK
 SKFSTADEARONPVPKAMEVYVNGSGEWSNGGLLVIDASLEICRLEPYAPS
 NVTSVCSSEGESEVWCLDKANSLVYHSTYQOLCARFCVSPSLDMFPYAPL
 DTEPPAASHTANPKYEDSIVADVISIMSELTQILIHQESLIDYCSMSYSPPR
 QAARSPSLSPASSSVPEFSDCEDSDMLHTPGAASDRSHDLTPMGFTESQHO
 AKTILAVRDLIWPVRGDIVYIGLEKSENORGVIAVTKARELTPHGVYDAVAVA
 KDTVYCTEENETENGLAVMRGMRAREDDIFQSTIEELERLEACTRKRK"

Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-836-392-21 (1-728) x AB058693 (1-5370)
 QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheserGluPheArgGlnGlu 20
 DB 1926 ATGGCTAGGACCTGGGGCCACCGATGCCAAGAACTTCTCCGAGTCCGGAGGAG 1985
 QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
 DB 1986 GCCAGCATCTGCACCGCTGCAGACACCCCTGCATCTGCGCTCATCGGCATCGATC 2045
 QY 41 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 DB 2046 CACCGCGCTGCTGCGCCCTGGAGTGGCGCGCTGACGACGAGCTCAACACGCTGCTC 2105
 QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 80
 DB 2106 GAGAACGCCAGATCTTCTTATACCCCTGGGACACATCTCACCCAAAATATAGCC 2165
 QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
 DB 2166 TACCAATGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2225
 QY 101 LysSerAspAsnIleLeuValTyrSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
 DB 2226 AGTCGGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2285
 QY 121 SerAspTyrGlyIleSerArgLysPheHisGlnGlyAlaLeuGlnLysIleGlyThr 140
 DB 2286 TCTGATACGAGGATTCGAGGACATCATTCATGAGGCGCCCTGAGGCGGCGACT 2345
 QY 141 ProGlyTyrGlnAlaProGluLeuArgProAlaGlyIleValTyrAspGlnLysValAspMet 160
 DB 2346 CCTGCTACACAGCCGCCAAGATCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2405
 QY 161 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHis 180
 DB 2406 TCTCTCTAGGAATGGTGGTCTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2465
 QY 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnPro 200
 DB 2466 CACCACTCCAGATTCACAAAGAGCTGCTCAAGGCGATCGCCGCTGCTGCTGCTGCTG 2525
 QY 201 GluGlnValGlnPheArgArgLeuGlnAlaLeuMetMetGlnCysTrpAspThrLysPro 220
 DB 2526 GAGGAAGTCAGTTCGGGCGACTGCGAGGCTCATGATGAGTGTGGGACACTAAGCCA 2585
 QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
 DB 2586 GAGAACGACCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2645
 QY 241 PheMetTyrGlnLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlnGln 260
 DB 2646 TTTCTATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2705
 QY 261 TyrThrValValPheThrAspGlyLysGlnGlnLysSerArgAsnTyrThrValValAsnThr 280
 DB 2706 TACACCGTGGTGTGGATGATGAAGAGGAGTCCAGGACTACCGGTGTGTAACACA 2765
 QY 281 GluLysGlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
 DB 2766 GAGAACGCGCTCATGAGAGTGCAGAGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2825
 QY 301 LeuGlnValGlnArgSerLeuThrPheAlaThrGlnAspGlnLysIleTyrIleTyrThr 320
 DB 2826 CTCACAGTCCAGAGATCCCTGTGACAGCAGCCAGGACCAAGAAATCATCATCAACCC 2885
 QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
 DB 2886 CTCAGAGGATGTGCTTAAACACACCCCAAGGCGCTGTGATATCTCCAGCTGTGCTG 2945

BASE COUNT 1226 a 1553 c 1502 g 1089 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,36e-270
 Score: 3822.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%

Length: 5370
 Matches: 728
 Conservative: 0
 Mismatches: 0

QY 341 ThrCysPheLeuAlaValProValIleTylSAsnSerTyrLeuValLeuAlaGlyLeu 360
 |||||
 Db 2946 ACCGTGTTTGGCCCTGGCTGTATTAAAAAAGATTCCTACTGCTTACGGGGCTTC 3005
 QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 |||||
 Db 3006 GCCGATGGCTGTGTGCTGTGTCCCGTGGTGGGACACCCAAAGACAGCTCTCC 3065
 QY 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspLysAlaArg 400
 |||||
 Db 3066 TACCTGTGTACACACACACACCAAGCTCCAGATTGAGATCGCGATGAAGACGACGG 3125
 QY 401 GluAsnProTyrProValLysAlaMetGluValAlaAsnSerGlySerGluValTyrPyr 420
 |||||
 Db 3126 CAGAACCCCTACCCAGTGAAGGCAATGAGGTGTGAACAGCGCTCTGAGGCTGTGAC 3185
 QY 421 SerAsnGlyProGlyLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
 |||||
 Db 3186 AGCAATGGGCGGCGCTCTGTGATGATGAGTGTCTCCCTGGAGATGTGACAGGGGCTG 3245
 QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGluArgGly 460
 |||||
 Db 3246 GAGCCCTACATGGCCCCCTCCATGATGATGATGATGATGATGATGATGATGATGATG 3305
 QY 461 GluGluValValTyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 |||||
 Db 3306 GAGGAGT 3365
 QY 481 ThrTyrGluLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
 |||||
 Db 3366 ACCTACAGCTGT 3425
 QY 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
 |||||
 Db 3426 CCGCTGCGGCGCTGTGACACGAGAACCCCGGACGACGACGACGACGACGACGACG 3485
 QY 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluGluLeuGlyThrGln 540
 |||||
 Db 3486 CCGAGAGGGAGCTCCATCGCGAGCGTGAAGATATGATGATGATGATGATGATGATG 3545
 QY 541 IleLeuIleHisGlnGluSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 560
 |||||
 Db 3546 ATCTGTATCCACACGAGATCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3605
 QY 561 ProProArgGlnAlaAlaAspSerProSerSerLeuProSerSerProAlaSerSerSer 580
 |||||
 Db 3606 CCACCCCGGACGCTGCGAGGTGCTCCCTCAAGCTCCCTCCACGACGATTTCTTCC 3665
 QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
 |||||
 Db 3666 AGTGTGCTTTCTCCACGCTGCGAGGACTGACAGATGCTTACATACGCGGCTGCTCC 3725
 QY 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
 |||||
 Db 3726 TCCGACAGGCTGAGCATGACATGACCCCATGAGAGGGGAGACCTTCAACGACGACCTG 3785
 QY 621 GluAlaValLysIleLeuAlaValArgAspLeuIleTyrValProArgArgGlyLysPhe 640
 |||||
 Db 3786 CAGGCGGTGAAGATCCCGCGCTGAGACCTATTTGGGTCCCGCGGCGGAGAT 3845
 QY 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 |||||
 Db 3846 GTTATGCTCATGTGCTGCGGAGAAAGATTGAGAGCCGAGGCGGCGAGTCATTGCTGCTC 3905
 QY 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 680
 |||||
 Db 3906 TTTAAAGCCGAGAGGATGATCCGCAATGGGGTCTGTGTGATGTGCGGAGTGGCAAG 3965
 QY 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrCysLeuAlaValTyrArg 700
 |||||
 Db 3966 GACACTGTGTGTGACCTTTGAAATGAAGAAACACAGAGTGTGCTGCGCGCTGTGAGG 4025
 QY 701 GlyTyrGlyAlaArgGluPheAspIlePheTyrGlnSerTyrGluGluLeuArgLeu 720

Db 4026 GGCTGGGGGCGGACGAGATGTGACATTTCTACCACTCTACGAGAGCTGGCGGGCTG 4085
 QY 721 GluAlaCysThrArgLysArgArg 728
 |||||
 Db 4086 GAGGCTGTGCTGCGCAAGAGAGG 4109
 RESULT 2
 AK058028
 LOCUS AK058028 4210 bp mRNA linear PRI 31-OCT-2001
 DEFINITION Homo sapiens cDNA FLJ25299 f1s, clone STW07710.
 ACCESSION AK058028
 VERSION AK058028.1 GI:16554029
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens stomach mucosa cDNA to mRNA, clone_1lb:STW
 clone:STW07710.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS 1
 Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,
 Ono, Y., Hotuka, T., Hirooka, S., Murakawa, K., Takiguchi, S.,
 Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,
 Yamashita, H., Chiba, Y., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S.,
 Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
 Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and
 Sugano, S.
 NEDO human cDNA sequencing project
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 4210)
 AUTHORS Sugano, S. and Suzuki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan
 (E-mail:cdna@igms.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan. cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction and 5'-end one pass sequencing: Institute of Medical
 Science, University of Tokyo, Laboratory of Genome Structure, Human
 Genome Center; 3'-end one pass sequencing: RAB; clone selection for
 full insert sequencing: RAB and Helix Research Institute.
 FEATURES
 source
 1..4210
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="STW07710"
 /tissue_type="stomach mucosa"
 /clone_1lb="STW"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 976 a 1198 c 1157 g -- 879 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,54e-265 Length: 4210
 Score: 3811.00 Matches: 726
 Percent Similarity: 99.73% Conservative: 0
 Best Local Similarity: 99.73% Mismatches: 2
 Query Match: 99.71% Indels: 0
 DB: 9 Gaps: 0
 US-09-836-392-21 (1-728) x AK058028 (1-4210)
 QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 20
 |||||
 Db 769 ATGCTGAGGACCTCGCGGCGACCATGCGATGAAGACTTCCGAGTTCGCGGAGAG 828
 QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
 |||||
 Db 829 GCCAGATGCTGACCGCGCTGACGACACCCCTGCATGCTGCGCTACCGCATCAGCATC 888

```

OY 41 HisProLeucCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
Db 889 CACCCGCTCTCTGCGCTGAGAGTGGCGCGCTCAGACCGCTCAACACCGTGTGTC 948
OY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnLysIleAla 80
Db 949 GAGAACGGCAGAGATTCTCTTATACCCCTGGACACATGCTCACCCCAAAATATACC 1008
OY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
Db 1009 TACACAGATCGCCCTGGCGCTGCTACCTGCACAAAGAAAACATCATCTCTGTGACCTA 1068
OY 101 LysSerAspAsnIleLeuValIlePheSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
Db 1069 AAGTCGGACACATTCGTGTGTGTGCTTACGTCAGGAGGACATCAATCAATCAAGCTA 1128
OY 121 SerAspTyrGlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnLysThr 140
Db 1129 TCTGACTACGGGATTTCCAGCAGCATTCATGAGGCGCCCTAGGCGTGGAGGCGACT 1188
OY 141 ProGlyTyrGlnAlaProGlyIleArgProArgIleValTyrAspGlyLysValAspMet 160
Db 1189 CCTGGCTACAGCGCCCGCAGATCAGCGCTGCATGTATATGATGAGAGATAGATATG 1248
OY 161 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlyHis 180
Db 1249 TTCCTCATATGGAATGTGCTTACAGATGTGTGTGTCAGAGACGCGCTGACCTGGCCAC 1308
OY 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnPro 200
Db 1309 CACCACTCCAGATTCGCAAAAGCTGTCCAAAGGACATCGCGCTGTGGGAGAGCG 1368
OY 201 GluGlnValGlnPheArgArgLeuGlnAlaLeuMetMetGlyCysTyrAspThrLysPro 220
Db 1369 GAGAAAGTCAAGTTCGCGCGCTGCGAGCTGCATGATGAGTGTGGAGACCTAAGCA 1428
OY 221 GluLysArgProLeuAlaLeuSerValLysSerGlnMetLysAspProThrPheAlaThr 240
Db 1429 GAGAAAGTCAAGTTCGCGCGCTGCGAGCTGCATGATGAGTGTGGAGACCTAAGCA 1488
OY 241 PheMetTyrGlnLeuCysGlyGlyGlnThrAlaPhePheSerSerGlnGlyGlnGlu 260
Db 1489 TTCATGTATGAACGTGTGTGTGGAGACAGACAGCTTCTCTCATCCAGGCGCAGAG 1548
OY 261 TyrThrValValPheTyrAspGlyLysGlnLeuSerArgAsnTyrThrValValAsnThr 280
Db 1549 TACACCGTGTGTGTGGAGTGAAGAAAGAGATCCAGAACTACACGCGTGTGAACACA 1608
OY 281 GluLysGlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
Db 1609 GAGAAAGTCAAGTTCGCGCGCTGCGAGCTGCATGATGAGTGTGGAGACCTAAGCA 1668
OY 301 LeuGlnValGlnArgSerLeuTyrPheAlaThrGlnAspGlnLysIleTyrIleTyrThr 320
Db 1669 CTCAGAGTCCAGAGATCCCTGTGGACACCAACGAGACAGAAATATCATCATCACACC 1728
OY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
Db 1729 CTCAGAGGATGTCCTCTTAAACACACCCCAACAGCGCTTGATCTCCAGCTCTGCTC 1788
OY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
Db 1789 ACCTGCTCTTGGCGCGCTGCTGTATTAAGAAATCTCTACCTGTGTGTGGCGGCTC 1848
OY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
Db 1849 GCCGATGGCTTGTGGCTGTGTTCCTGCTGTGGCGGACCCCAAGAGACAGCTGCC 1908
OY 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGlnAspAlaArg 400
Db 1909 TACCTGTGCTCACACACAGCCCAAGGTCAGAGTTCACATCGCGGATGAAGACCCAGG 1968
OY 401 GlnAsnProTyrProValLysAlaMetGlnValValAsnSerGlySerGlnValTyrPyr 420

```

```

Db 1969 CAGAAACCCCTACCGAGTGAAGGCCATGAGGTGTCAACAGCGCTCTGAGGTGTGATC 2028
OY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnLysCysArgArgLeu 440
Db 2029 AGCAATGGCGCGGCGCTCTCTGTATGACAGCTGTGCTCCCGAGATCTGACGCGGCTG 2088
OY 441 GluProTyrMetAlaProSerMetValIleSerValValCysSerSerGlyGlyArgGly 460
Db 2089 GAGCCCTACATGAGCGCCCTCCATGTTACGTACATGCTGTGTGGCTGTGAGGAGAGAGG 2148
OY 461 GluGlnValValIleTyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
Db 2149 GAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2208
OY 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValAlaProSerProLeuArgAspMetPhe 500
Db 2209 ACCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2268
OY 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
Db 2269 CCCGTGGCGCGCTGTGAGACAGGAACCCCGGACAGCCACAGGCCCAACCAAGGCTG 2328
OY 521 ProGlnLysAspSerIleAlaAspValSerIleMetTyrSerGlnGlnLeuGlyThrGln 540
Db 2329 CCTAGGGGGACTCATCGCGAGTGCAGCATCATGATGATGAGAGCTGGGACAGCAG 2388
OY 541 IleLeuIleHisGlnGlnLeuSerLeuThrAspTyrCysSerMetSerSerTyrSerSer 560
Db 2389 ATCTGATCCACAGGATCTCATCTGATCTGCTCATGCTCTCTCTCTCTCTCTCTCTCT 2448
OY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
Db 2449 CCACCGCGCGAGCTGTGAGAGTCCCGCTCAAGCTCCCGACAGCTCCCGCAAGTCTTCC 2508
OY 581 SerValProPheSerThrAspCysGlnAspSerAspMetLeuHisThrProGlyValAla 600
Db 2509 AGTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2568
OY 601 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeu 620
Db 2569 TCCGACAGTGTGAGCATGACATGACCCCATGAGAGGAGAGACTTACGACACAGCTG 2628
OY 621 GlnAlaValLysIleLeuAlaValAlaArgAspLeuIleTyrValProArgArgGlyAsp 640
Db 2629 CAGCGCGTGAAGATCTCTGCGGTGAGAGCTCATTTGGGTGCTCCAGCGCGGTGAGAT 2688
OY 641 ValIleValIleGlyLeuGlnLysAspSerGlnLysGlnArgGlyValArgValIleAlaVal 660
Db 2689 GTTATCGTATGCTGTGAGAGATGATCTGCGCGCCAGCGGCGCGAGTATTTGCCGTG 2748
OY 661 LeuLysAlaArgGlnLeuThrProHisGlyValLeuValAlaAspAlaValAlaLys 680
Db 2749 TTTAAAGCCCGAGAGCTGATCCCATGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2808
OY 681 AspThrValValCysThrPheGlnAsnGlnAsnThrGlnTyrCysLeuAlaValAlaTyrArg 700
Db 2809 GACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2868
OY 701 GlyTyrGlyValAlaArgGlnPheAspIlePheTyrGlnSerTyrGlnGlnLysArgLeu 720
Db 2869 GGCTGGGGCGCGAGGAGTGTGACATTTTCTTACCAATCTTACGAGAGACTGGGCGGCTG 2928
OY 721 GluAlaCysThrArgLysArgArg 728
Db 2929 GAGGCTTCACTCCAGAGAGAGG 2952

```

RESULT 3
 AX250161
 LOCUS AX250161 6045 bp DNA linear PAT 28-SEP-2001
 DEFINITION Sequence 6 from Patent WO016594.
 ACCESSION AX250161
 VERSION AX250161.1 GI:15864536

KEYWORDS human
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 6045)
AUTHORS Plowman, G.D., Whyte, D., Manning, G., Sudarsanam, S. and Martinez, R.
TITLE Human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 016594-A 6 13-SEP-2001;
Sugen, Inc. (US)
FEATURES
source Location/Qualifiers
1..6045
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1359 a 1779 c 1707 g 1200 t
ORIGIN
Alignment Scores:
Pred. No.: 2,33e-252 Length: 6045
Score: 3579.50 Matches: 688
Percent Similarity: 94.518 Conservative: 0
Best Local Similarity: 94.518 Mismatches: 1
Query Match: 93.668 Indels: 39
Gaps: 1
US-09-836-392-21 (1-728) x AX250161 (1-6045)
QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheserGluPheArgGlnGlu 20
DB 3976 ATGCTAGGAGCACTGGCGGCGACCGATGCCATGAGAACTTCCGAGTTCCGGCAGGAG 4035
QY 21 AlSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
DB 4036 GCCAGATGCTGCACCGCGCGACGCCCTGCATCGTGGCGCTCATCGGCATCAGATC 4095
QY 41 HisProLeuGlySerPheAlaLeuGlnLysAlaProLeuSerSerLeuAsnThrValIleSer 60
DB 4096 CACCCGCTGCTTCCGCCCGCGAGCTCGCCGCTCAGCAGCTCAACACCGCTGCTCC 4155
QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 80
DB 4156 GAGAGCGCCAGAGATCTTCTTATACCCCTGGGACACATGCTCAACAAAAATAGCC 4215
QY 81 TyrGlnIleLeuAlaSerGlyLeuAlaTyrLeuHisLysAsnIleIlePheCysAspLeu 100
DB 4216 TACCAATCGCTCGCGCGCTGCTACCTGCACAAAGAAATCATCTTCTGTGACCTG 4275
QY 101 LysSerAspAsnIleLeuValTyrSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
DB 4276 AAGTCGACACACTTCTGGTGTGCTCCCTTGACGTCAAGAGCACATCAACATCAAGCTA 4335
QY 121 SerAspTyrGlyIleSerArgLysSerPheHisGlnGlyAlaLeuGlnLysIleThr 140
DB 4336 TCTGACATACGAGATTCGAGGAGCTATCTCATGAGGCGCCCTAGGCGTGGAGGACT 4395
QY 141 ProGlyTyrGlnAlaProGlyIleArgProAlaGlyIleValTyrAspGlyLysValAspMet 160
DB 4396 CCGGCTACACAGGCCCGCGAGATCAGGCTCGCATGTATATGATGAGAGGATGATGATG 4455
QY 161 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHis 180
DB 4456 TCTCTCATGGAATGGGCTCTAGCACTTCTGTCAAGACAGCGCCCTGCACAGTGGCCAC 4515
QY 181 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnLysPro 200
DB 4516 CACCAAGCTCCAGATTCGCAAGAGCTGTCCAAAGGCAATCGCGGTTCTGGGCGACCCG 4575
QY 201 GluGlnValGlnPheArgGlnGlnAlaLeuMetMetLysCysTrpAspThrLysPro 220
DB 4576 GAGAGATGCAAGTTCGCGCACTGCAAGCGCTCATGATGAGTGTGGACACTAAACCA 4635
QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240

DB 4636 GAGAAACGACCGCTGCGCTGTGAGTGTGAGCAAGATGAGACCCGACTTTTGGCCACC 4695
QY 241 PheMetTyrGlnLeuGlySerGlyLysGlnThrAlaPhePheSerSerGlnGlnGlu 260
DB 4696 TTCATGTATGAACTGTGCTGGGAGGAGACAGCCCTTCTTCCATCCAGCGCCGAGGAG 4755
QY 261 TyrThrValValPheTrpAspGlyLysGlnLysSerArgAsnTyrThrValValAsnThr 280
DB 4756 TACCCGCTGTGTGTTGGATGTGAAAGAGAGCTCCAGAACTACAGCTGTGGAACACA 4815
QY 281 GluLysGlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
DB 4816 GAGAAAGGCGCTCATGAGAGTGCAGAGAGATGTGCTGCTGGATGAGAGTGCAGCTCCAG 4875
QY 301 LeuGlnValGlnArgSerLeuThrPheAlaThrGlnLysGlnLysIleTyrIleThr 320
DB 4876 CTCGAGGTCCAGATCCCTGTGAGACAGCCAGCAG----- 4911
QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAla 340
DB 4911 ----- 4911
QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
DB 4912 ----- 4912
QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
DB 4939 GCCGATGGGCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4998
QY 381 TyrLeuGlySerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
DB 4999 TACCTGTCTCACACACAGACAGACAGAGTCCAACTTCCAGATCCGATGAGACGCAAG 5058
QY 401 GlnAsnProTyrProValLysAlaMetGlnValAlaAsnSerGlySerGlnValTyrTyr 420
DB 5059 CAGAACCCCTACCACTGAGGAGGATGAGAGGCTGCAACAGCGGCTGAGGCTGCTGCTGCT 5118
QY 421 SerAsnGlyProGlyLeuValIleAspCysAlaSerLeuGlnIleCysArgAlaGln 440
DB 5119 AGCAATGGGCGGCTCTCTGTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5178
QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnLysArgGly 460
DB 5179 GAGCCCTACTGAGCCCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 5238
QY 461 GluGlnValValTyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
DB 5239 GAGAGAGTGTGTGTGCTGATGAGCAAGGCCAATCTCTGTGTGATGATGATGATGATGATG 5298
QY 481 ThrTyrGlnLeuGlyAlaArgTyrPheCysGlyValAlaProSerProLeuArgAspMetPhe 500
DB 5299 ACCTACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5358
QY 501 ProValArgProLeuAspThrGlnProProAlaAlaSerHisThrAlaAsnProLysVal 520
DB 5359 CCCGTGGGCGCTTGGACAGGAGAACCCCGCGACGCCACACAGGCCAACCCAAAGAGTG 5418
QY 521 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGlnLysGln 540
DB 5419 CTTGAGAGGAGACATCCAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATG 5478
QY 541 IleLeuIleHisGlnLysSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 560
DB 5479 ATCTGTATCCACAGGAAATCACTGATGATGATGATGATGATGATGATGATGATGATG 5538
QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
DB 5539 CCACCCCGGACAGGCTGCGAGAGTCCCTCAAGCTCCCAAGCTCCCAAGCAATTTCTCC 5598
QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
DB 5599 AGTGTGCTTCTTCCACCGAGCTCGAGAGACTCAGACATGCTACATACGCGCGGTGCTGCC 5658

QY 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
 DB 5659 TCCGACAGCTGATGACATGACCTGACCCCATGAGCGGAGACCTTCAGCCAGCACCTG 5718
 QY 621 GlnAlaValAlaLeuLeuAlaValaAspAspLeuIleTyrValProArgArgGlyGlyAsp 640
 DB 5719 CAGCCCGTAAAGATCTCTCCCGCTCAGACACCTCATTTGGGTCCCGCGCGGTGAGAT 5778
 QY 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 DB 5779 GTATGCTATGCTGCTGAGAAAGATTTCTGGCGCCAGCGGCGCCGAGTCATGCGCTG 5838
 QY 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValaAspAlaAlaValAlaLys 680
 DB 5839 TTTAAAGCCCGAGAGCTGACCTCGGAGGTGCTGTGGATGCTGCGGTGGCGAAG 5898
 QY 681 AspThrValValCysThrPheGluAspGluAsnThrGluTyrCysLeuAlaValTyrArg 700
 DB 5899 GACACTGTTGTGTGACCTTTGAAATGAAACACAGAGTGTGTGCTGCGCTGTGAGAG 5958
 QY 701 GlyTyrGlyAlaArgGluPheAspIlePheTyrGlnSerTyrGluGluLeuGlyArgLeu 720
 DB 5959 GGCTGGGGGCGCAGAGAGTTCACATTTTCTACCACTCTACAGAGAGCTGGCGCGCTG 6018
 QY 721 GluAlaCysThrArgLysArgArg 728
 DB 6019 GAGCCTTGCACTGCGAAGAGAG 6042
 RESULT 4
 LOCUS BC005408 3509 bp mRNA linear PRI 12-JUL-2001
 DEFINITION Homo sapiens, similar to hypothetical protein FLJ23119, clone
 MG:5630 IMAGE:3463078, mRNA, complete cds.
 ACCESSION BC005408
 VERSION BC005408.1 GI:13529310
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 3509)
 Strausberg, R.
 Direct Submission
 Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Sequencing Group at the Stanford Human
 Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickson@stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
 Series: IRAK Plate: 4 Row: n Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 FEATURES
 source
 1. 3509
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:5630 IMAGE:3463078"
 /tissue_type="Cervix, carcinoma"

CDS
 /clone_11b="NIH_MGC_12"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1317..1982
 /codon_start=1
 /product="Similar to hypothetical protein FLJ23119"
 /protein_id="AAH05408.1"
 /db_xref="GI:13529311"
 /translation="MYSTTQOLCARHYRCGYSPRLDMFPYRPLDTPPAASHANK
 VPEGDSIDYVIMSEELGTLHSELDYICSSSTSSPPQAKRSPLSPFA
 SSSVPESTDSDMLMPPAASDRSHDLTPMDETFSDHDAVALIARLDLWVP
 RRGDVIVIGLEKSEARGRVIAVLKARELTPHGIIMPVSSVKWAGWPRDVTYMA
 AVM"
 BASE COUNT 810 a 994 c 953 g 752 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.31e-252 Length: 3509
 Score: 3578.50 Matches: 690
 Percent Similarity: 93.12% Conservative: 0
 Best Local Similarity: 93.12% Mismatches: 1
 Query Match: 93.63% Indels: 50
 DB: 9 Gaps: 1
 US-09-836-392-21 (1-728) x BC005408 (1-3509)
 QY 38 IIEserIIEHisProLeuGluPheAlaLeuGluAlaProLeuSerSerLeuAsnThr 57
 DB 3 ATCAGCATCCACCCGCTGCTGCTGCGCTGGAGCTGGCCGCTCAAGCCTCAACACC 62
 QY 58 ValLeuSerGluAsnAlaArgAspSerPheIleProLeuGlyHisMetLeuThrGln 77
 DB 63 GTCGCTGCGAAGAGCCAGAAATCTCTTATACCCCTGGACATGCTACCCAA 122
 QY 78 LysIleAlaTyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePhe 97
 DB 123 AAAATAGCTACAGATCCCTCGGGCGCTGCTACCTGCACAAAGAAACATCATCTTC 182
 QY 98 CysAspLeuLysSerAspAsnIleLeuValTyrPserLysAspValLysGlnHisIleAsn 117
 DB 183 TGTGACCTGAACTCGACAAACATTTCTGTGTGTGCTGCTGACGTCAAGAGCATCAAC 242
 QY 118 IleLysLeuSerAspTyrGlyIleSerArgIleSerPheHisGlyLysAlaLeuGlyVal 137
 DB 243 ATCAAGCATATGCTAGTACAGGATTTTCAGAGCATATTCATCAAGAGCGCCCTAGCGCTG 302
 QY 138 GluGlyThrProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLys 157
 DB 303 GAGGCGACCTCCGTGCTACAGGCGCCAGAGATCAGGCTCGCATGTATATGATGAGAAG 362
 QY 158 ValAspMetPheSerTyrGlyMetValLeuTyrGluLeuLeuSerGlyGlnArgProAla 177
 DB 363 GTAGATATGTTCTCTATGGAATGTGTCTTACGAGTTGCTGCGACAGCGCCCTCA 422
 QY 178 LeuGlyHisHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeu 197
 DB 423 CTGGGCGAACCACTCAAGTTCACAAAGAGCTGTCCAAAGGATCCCGCGCTTCTG 482
 QY 198 GlyGlnProGluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGluCysTyrAsp 217
 DB 483 GGGCAGCGCGAGGAGATGCACTCCGCGGACGACGCGGCTCATATGAGAGTGGGAC 542
 QY 218 ThrLysProGluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThr 237
 DB 543 ACTAAGCCAGAAAGCAGCCGCTGCTGTGGGTGGAGGACCAATAGAGACCGGACT 602
 QY 238 PheAlaThrPheMetTyrGluLeuCysGlyGlyGlnThrAlaPhePheSerSerGln 257
 DB 603 TTTCGACACTCATATGAACTGCTGTGGAAAGACAGACGCTTTCTCATCCAG 662
 QY 258 GlyGlnGluTyrThrValValPheThrAspGlyLysGluGluSerArgAsnTyrThrVal 277
 DB 663 GGCAGAGATACACCGTGTGTGTTTGGATGGAAGAGAGTCCAGAACTACAGCGGTG 722

/clone="BNGH41000104"
 /cell_line="H4"
 /cell_type="neuroglioma"
 /clone_id="BNGH41"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 707 a 872 c 862 g 671 t
 ORIGIN

Alignment Scores:

Pred. No.: 8,94e-227 Length: 3112
 Score: 3225.00 Matches: 616
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 84.38% Indels: 1
 DB: 9 Gaps: 0

US-09-836-392-21 (1-728) x AK054662 (T-3112)

Oy 112 ValysgluHs11easnlleusSeraspTyrglyIleSerArgGlnSerPheHs 131
 Db 2 GTCAGAGGACATCAATCACTATCTACTACGAGGATTCGAGCAGCTCATTCAT 61
 Oy 132 GluGluAlaLeuGlyValGluGlyThrProGlyTyrGlnAlaProGlnIleArgProArg 151
 Db 62 GAGGCGCCCTAGCGGTGGAGGCACTCTGGCTACAGGCCCCAGAGATCAGGCTCGC 121
 Oy 152 IleValTyrAspGluTyrValAspMetPheSerTyrGlyMetValLeuTyrGluLeu 171
 Db 122 ATTGTATATGATGAGAGGTAGATATGTCTCCATGAGAAAGTGCTCTAGAGTGTG 181
 Oy 172 SerGlyGlnArgProAlaLeuGlyHs11easnlleuGlnIleAlaLysLysLeuSerLys 191
 Db 182 TCAGGACAGCCCTGCGCTGGCCACACAGCTCCAGATGCCAAGAGCTGCCAAG 241
 Oy 192 GlyIleArgProValLeuGlyGlnProGluGluValGlnPheArgArgGlnAlaLeu 211
 Db 242 GGCATCCGCGCGGTCTGGGCGAGCGGAGAGAGTGCGAGTCCGCGCTCAGCGCTC 301
 Oy 212 MethetGluCysTrpAspThrLysProGluTyrAspProAlaLeuSerValAlaSer 231
 Db 302 ATGATGAGTCTGGGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
 Oy 232 GlnMetLysAspProThrPheAlaThrPheMetTyrGluLeuGlyLysGlnThr 251
 Db 362 CAGATGAG 421
 Oy 252 AlaPhePheSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 271
 Db 422 GCGTCTCTCATCCAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
 Oy 272 SerArgAsnTyrThrValAlaAsnThrGluLysGlyLeuMetGluValAlaArgMetCys 291
 Db 482 TCCAGGAACACAGCGTGTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
 Oy 292 CysProGlyMetLysValSerCysGlnLeuGlnValGlnArgSerLeuTyrThrAlaThr 311
 Db 542 TGCCTGGAG 601
 Oy 312 GluAspGlnLysIleTyrIleTyrThrLeuLysGlyMetCysProLeuAsnThrProLys 331
 Db 602 GAGGACGAGAAATCTATCATCTACACCTCAAGGAGCATGTCCTTAAACACACCCAA 661
 Oy 332 GlnAlaLeuAspThrProAlaValAlaThrCysPheLeuAlaValProValIleLysLys 351
 Db 662 CAGGCTTGATGATCCAGCTGTGTCTACCTGCTTCTGGCCGCTGCTTATTTAAAG 721
 Oy 352 AsnSerTyrLeuValLeuAlaGlyLeuAlaAspGlyLeuValAlaValPheProValAla 371
 Db 722 AATTCCTACGTGCTTACGAGGCGCTCGCATGAGGCTTGAGGCTGTGCTTCCCTGCTG 781
 Oy 372 ArgGlyThrProLysAspSerCysSerTyrLeuGlySerHisThrAlaAsnArgSerLys 391
 Db 782 CGGGGACACCCAAAGAGAGAGCTGCTCTTACTGTGCTCACACAGCCAAAGAGGTCCAAG 841

Oy 392 PheSerIleAlaAspGluAspAlaArgGlnAsnProTyrProValLysAlaMetGluVal 411
 Db 842 TTCAGATCGGGGATGAAAGAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
 Oy 412 ValAsnSerGlySerGluValIleTyrTyrSerAsnGlyProGlyLeuValIleAspCys 431
 Db 902 GTCACAGCGGCTCTGAGGTCTGTGTACAGCAATGGGCGGCGCTTGTGATCGACTGT 961
 Oy 432 AlaSerLeuGlnIleCysArgArgLeuGlnProTyrMetAlaProSerMetValThrSer 451
 Db 962 GCTTCCCTGGAGATCTGAGGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
 Oy 452 ValValCysSerSerGluGlyArgGlyGluGluValAlaTyrCysLeuAspLysAla 471
 Db 1022 GTGCTGTGAGCTGTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
 Oy 472 AsnSerLeuValMetTyrHisSerThrTyrGlnLeuGlyAlaArgTyrPheCysGly 491
 Db 1082 AACTCTTGGTGTATGATACCATCCACACCTACAGCTGTGTGCGCTTCTGCGGG 1141
 Oy 492 ValProSerProLeuArgAspMetPheProValArgProLeuAspThrGluProAla 511
 Db 1142 GTCCCAAGCCCTCAGAGGACATGTTCCGTGCGGCTTGGACAGAGAGAGAGAGAG 1201
 Oy 512 AlaSerHisThrAlaAsnProLysValProGluLysSerIleAlaAspValSerIle 531
 Db 1202 GCCAGCCACAGCGGCAACCCAAAGTGCTGAGGAGGAGAGAGAGAGAGAGAGAGAG 1261
 Oy 532 MetTyrSerGluGluLeuGlyThrGlnIleLeuIleHisGlnGlnSerLeuThrAspTyr 551
 Db 1262 ATGTACAGTGTAGGAGCTGGGCGACAGATCTGTATCCACAGAGAAATCACTACAGCTAC 1321
 Oy 552 CysSerMetSerTyrTyrSerSerSerProProArgGlnAlaAlaArgSerProSerSer 571
 Db 1322 TCTTCCATGCTCTCTTACTCTCTTATCCCAAGCCCGGAGAGGTGCCAGAGTCCCTCAAGC 1381
 Oy 572 LeuProSerSerProAlaSerSerSerSerValProPheSerThrAspCysGluAspSer 591
 Db 1382 CTCGCCAGCTCCAGAGAGATTTCCAGAGTGTGCTTCTTCCACAGCTGAGAGAGCTCA 1441
 Oy 592 AspMetLeuHisThrProGlyAlaAlaSerAspArgSerGlnHisAspLeuThrPromet 611
 Db 1442 GACATGTCTACATACGCGCGGTGCTGCTCCAGAGGTCTGTGAGTACCTGACCCCATG 1501
 Oy 612 AspGlyGluThrPheSerGlnHisLeuGlnAlaValLysIleAlaValAlaArgAspLeu 631
 Db 1502 GACGGGAGACTTCAGACACACCTGAGGCGGTGAAGATCTGCGCGTACAGAGACTTC 1561
 Oy 632 IleTyrValProArgArgGlyGlyAspValIleValIleGlyLeuGluLysAspSerGlu 651
 Db 1562 ATTGGGTCGCCAGGCGGTGAGAGATGTATCTCATTTGCGCGGAGAGAGATTCGAA 1620
 Oy 652 AlaGlnArgGlyArgValIleAlaValLeuLysAlaArgGluLeuThrProHisGlyVal 671
 Db 1621 GCCCAGGCGGCGGAGATGATGCGGTCTTAAAGCCGAGAGCTGATCCGCAAGGGGTG 1680
 Oy 672 LeuValAspAlaAlaValAlaValLysAspThrValLysThrPheGluAsnGluAsn 691
 Db 1681 CTGGTGGATGCTGCGGTGGTGGCAAGAGACCTGTGTGTGACCTTGAAGAGAAAC 1740
 Oy 692 ThrGluTyrCysLeuAlaValAlaTyrArgGlyTyrPylAlaArgGluPheAspIlePheTyr 711
 Db 1741 ACAAGATGTGTGCGGCGGTGTGAGAGGCGTGGGCGGCGGAGAGAGATTCGATTTCTAC 1800
 Oy 712 GlnSerTyrGluGluLeuGlyArgLeuGluAlaCysThrArgLysArgArg 728
 Db 1801 CAGTCTACAGAGAGAGCTGGGCGGCTGAGAGCTTGACATCGCAAGAGAAAG 1851
 RESULT 6
 AK094556 2790 bp mRNA linear PRI 15-JUL-2002
 LOCUS Homo sapiens cDNA FLJ37237 f1s, clone BRAMY2002638.
 DEFINITION

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 1980)
Strausberg, R.
Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amgdbcm.tmc.edu
Gunnarane, P.H., Garcia, A.M., Lu, X., Huiyk, S.M., Hale, S.M.,
Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILN/ at: <http://image.llnl.gov>
Series: IRAC Plate: 37 Row: k Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES

source

Location/Qualifiers
1..1980
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:28646 IMAGE:4224922"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_1lb="NCI CGAP Co24"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
395..1153
/codon_start=1
/product="Similar to hypothetical protein FLJ23119"
/protein_id="AAH27199.1"
/db_xref="gi:20073185"
/translation="MYSATYQLCARPCGPNPDRFVSVPVLETPGSHKTTSG
FVEECIADVSNYSELETOILITDSDMSYSSSPHODPRSPSSISLTS
YSSVPSANYSRDSRLODESVTSDTEHDSIPMOETPSOHLAKRYVADLIMYPR
HGDIIVIGLEKDSQACRGRTIAYAKAEINRHGYLVAAVAAKDYVCGFRNEMTE
CLAVNRGAREFDIFVSYSELGRLACTRRR"

CDS

BASE COUNT 517 a 531 c 397 t
ORIGIN

Alignment Scores:

Pred. No.: 7e-112 Length: 1980
Score: 1653.50 Matches: 319
Percent Similarity: 87.07% Conservative: 17
Best Local Similarity: 83.07% Mismatches: 47
Query Match: 43.26% Indels: 1
DB: 10 Gaps: 1

US-09-836-392-21 (1-728) x BC027199 (1-1980)

QY 345 AAlaValProValIleLeuValSerIleValLeuAlaGlyLeuAlaAspGlyLeu 364
DB 2 GCAGTACCTGTTATCAAAAGAACTCCTCGTGTCTGCGAGCGCTGCTAGCGACAC 61
QY 365 ValAlaValPheProValValArgGlyThrProLysAspSerCysSerIleuCysSer 384
DB 62 GTGGCTGTCTTCTGCTGCGAGGGGACCCCAAAAGAAAGCTGCTCTACCTGTCTCTCC 121
QY 385 HIsThrAlaAsnArgSerLysPheSerIleAlaAspGlyAlaAlaArgGlnAsnProTyr 404
DB 122 CACACAGCCACAGGTCAGAGTTCTGCATCCCGACGAGAGATGACGAGAACCCCTTAC 181

QY 405 ProValLysAlaMetGluValAlaAsnSerGlySerGluValIlePyrSerAsnGlyPro 424
DB 182 CCAGTGAAGGCAATGGAAAGTGTCAAGTGGGTGTGAGTGTGTACAGCAGCGGCT 241
QY 425 GlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeuGluProTyrMet 444
DB 242 GGCTCTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
QY 445 AlaProSerMetValIleThrSerValValCysSerSerGluGlyArgGlyGluValAla 464
DB 302 GCTTCGCTCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
QY 465 TrpCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThrTyrGlnLeu 484
DB 362 TGGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
QY 485 CysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPheProValArgPro 504
DB 422 TGTGCGCGGTACTGTTGTGGGAGCCCAATCCTGACAGGAGACGTTTCTGTGAGCGCT 481
QY 505 LeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysValProGluGlyAsp 524
DB 482 TCAGTCTCAAGAAACCCCA---GGCAGTCACAAACACCCTCAAGGGCGCTGTGAGACAG 538
QY 525 SerIleAlaAspValSerIleMetTyrSerGluGluGlyThrGlnIleLeuIleHis 544
DB 538 TGCATCGCGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 598
QY 545 GlnGluSerLeuThrAspTyrCysSerMetSerTyrSerSerSerProArgGln 564
DB 599 CAGAGATGCTCAGCAGATGCTCAGTCTCTTATTTCTCTCCACCCACCCACCA 658
QY 565 AlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSerValProPhe 584
DB 659 GATCTAGTCCCTCCCTCAGTCTGCGACCTCCCTCAGCAGTATTCAGTGTGCTTTC 718
QY 585 SerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgSer 604
DB 719 TCTGCCACTATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
QY 605 GlnHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeuGlnAlaValLys 624
DB 779 GAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838
QY 625 IleLeuAlaValAlaAspLeuLeuThrValProArgArgGlyAspValIleValIle 644
DB 838 GTGCTTGTCTGTAAGAACTCATTTGGTCTTACATGAGGAGATATCATTTGTCTATC 898
QY 645 GlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaValLeuLysAlaArg 664
DB 899 GGCTGTGAGAAAGATTCAGGTGCTCCAGCGGGGACAGATCATGCTTTTAAACCCGA 958
QY 665 GluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLysAspThrValVal 684
DB 959 GAGCTCAATCGGACAGGGGTGTGTGTGTCGCCAGAGGTGGCAAGACACCTATGTG 1018
QY 685 CysThrPheGluAsnGluAsnThrGluThrProCysLeuAlaValIleArgGlyTyrGlyAla 704
DB 1019 TCGGCTTGTGCAATGAAACACACAGATGCTGCTGCTGTGAGAGGGGCTGGGGCGCC 1078
QY 705 ArgGluPheAspIlePheTyrGlnSerTyrGlnGluLeuGlyArgLeuGluAlaCysThr 724
DB 1079 AGGAGATTGATCATCTTACACAGTCATATGATGAGAGGTGGCGGCTGAGAGCTGTACT 1138
QY 725 ArgLysArgArg 728
DB 1139 CCAAGAGAGAGG 1150

RESULT 8
AL390202.01/c
WPCOMMENT
Sequence split into 10 fragments LOCUS AL390202 Accession AL390202
Fragment Name Begin End

Db	2713	ATCCCTGTCCTCTCTTTCAGAAATTCCTACTGCTGTATGCGGGCCCTCCCGATGGCGTT	2650
Qy	365	ValAlaValAlaPheProValValAlaGlyThrProLysAspSerCysSerTyrLeuCysSer	384
Db	2653	GTGGCTGTGTTTCCCGTGGCGGGGACCCCAAGACAGCGTCTCTCACTGTGCTCA	2594
Qy	385	H1sThrAlaAsnArgSerLysPheSerIleAlaAspGluAspIleArgLysAspProTyr	404
Db	2593	CACACACCAACAGAGTCCAAAGTTCAGCATGCGGGATGAAAGACGACGCAACCCCTAC	2530
Qy	405	ProValLysAlaMetGluValValAsnSerGlySerGluValTyrPyrSerAsnGlyPro	424
Db	2533	CCAGTGAAGGCGCAATGAGGTGTCAACAGCGGCTGTAGAGTGTGATACGAATGGCGCG	2470
Qy	425	GlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeuGluProTyrMet	444
Db	2473	GGCCCTCTTGTCATCGACTGTGGCTCCCTGGAATGTGACAGGGCGGTGAAGCCCTACATG	2410
Qy	445	AlaProSerMetValThrSerValValCysSerSerGluValArgGlyGluGluValAl	464
Db	2413	GCCCCCTCCATGTTAGTACGTACAGTCGTGTGAGCTTCGAGGGCAGAGGGAGAGGTGTCT	2350
Qy	465	TyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThrTyrGluLeu	484
Db	2353	TGTGCTCGTGTATGACAGGGCAACTCCTTGATGATGACACATCCACACCTACAGCTG	2290
Qy	485	CysAlaIleArgTyrPheCysGlyValProSerProLeuArgAspMetPheProValArgPro	504
Db	2293	TGTCCCGGATCTTCTGCGGGGTCGCCACGCCCTCAGGAGACATGTGTTCCCTGCGGGCCC	2230
Qy	505	LeuAspThrGluProProAlaAlaSerHisThrAlaAspProValValProGluGluAsp	524
Db	2233	TTGAGACACGGAAACCCCGGACGCCAGCAGCAGCAGCCCAACCAAGGTGCTTAGGGGGAC	2170
Qy	525	SerIleAlaAspValSerIleMetTyrSerGluGluLeuGluIleArgLysIleLeuIleHis	544
Db	2173	TCCATCCCGGACGTGACATCATGTACAGTGAAGAGACTGGGCACGCAGATCTGTATCCAC	2110
Qy	545	GluGluSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSerProProArgLys	564
Db	2113	CAGAAATCACTACTGTACTGTCTGTCTCATGCTCTCTACTCTCTCATCCACCCGCGCAG	2050
Qy	565	AlaAlaIleArgSerProSerSerLeuProSerSerProAlaSerSerSerValProPhe	584
Db	2053	GTCTCGCAGGTCCCGCTCAAGCCGCCAGCTCCCGACGAAAGTCTTCAGTGTGCTTTC	1990
Qy	585	SerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgSer	604
Db	1993	TCCACCCACTGTCCGGAGCTAGACATGTCTCAATACACCCCGGTCTCTCCGCAAGTCT	1930
Qy	605	GluHisAspLeuThrProMetLysArgLysIleThrPheSerGluHisLeuGluAlaValLys	624
Db	1933	GAGCATACCTGACCCCAATGACAGCGGGAGAACTTACAGCAGCACTGTGAGCGCGTAAG	1870
Qy	625	IleLeuAlaValArgAspLeuIleTyrValPro	635
Db	1873	ATCTGCTCCCGTACAGAGACGTCATTTGGGT-CCCAAGTACGTTTCCCGAGGTGAGGGCAC	1810
Qy	635		635
Db	1814	CATCCAGGACAGCCGCCACTGCTCTCTGTGGGACAGAGCAAGGGAGAGCCCGTGCT	1750
Qy	635		635
Db	1754	TCTTACAGGTCAATGTGGGAGCGAGGTGCTGTCTTCCAAAGAGGTTGTTTCTCAT	1690
Qy	635		635
Db	1694	TGGGAATGGAAGGAGCGGCTTTGACAAATCTAGAGTGGCGGACAAAGCTATTGTCT	1630
Qy	635		635

DB	1634	GCCAGAGAGCTTAAAGTGGCACACAGAGAAGCCACAGACTCAACTGACTGACTTCCAC	1575
OY	635	-----	635
DB	1574	TGAGAAATATCTCCCTCTGCTTTGAAGTCTGTCCCTCCCATTAATATCCCTTAGCAG	1515
OY	635	-----	635
DB	1514	AAATGTCATGTCGTGGCTCCCTCCCGACACTTCTCTGCACACCGGCTTCACTCCCGGA	1455
OY	636	-----	ArgArglygly 639
DB	1454	GAGCAGGGCTACCCAGCAATGCGCTTCCCTTCTGGGTTTGGTTCCTTAGGGCGCGTGA	1395
OY	640	AspValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAla	659
DB	1394	GATGTTATTCGTCATGTCGCTGGAGGAGGATTTCTGAAGCCAGCGGGCGGATCATTCGC	1335
OY	660	ValIleuLysAlaArgGluLeuThrProHisGly	670
DB	1334	GCTTAAAGCCCGAGAGCTGACTCGCATGGG	1302
RESULT 10			
AC090907/c			
LOCUS			
DEFINITION	AC090907	178529 bp	DNA linear HTG 01-JUN-2002
VERSION	AC090907		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.		
SOURCE	human		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 178529)		
JOURNAL	Birren,B., Linton,L., Nusbaum,C. and Lander,E.		
REFERENCE	unpublished		
AUTHORS	2 (bases 1 to 178529)		
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,		
	Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,		
	Camarata,J., Campopiano,A., Chang,Y., Choepel,Y., Colangelo,M.,		
	Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K.,		
	Diaz,J.S., Dodge,S., Ferro,S., Ferreira,P., Fitzhugh,W., Gage,D.,		
	Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,		
	Grand-Pierre,N., Hagos,B., Hesford,A., Horton,L., Hulme,W.,		
	Iliev,I., Johnson,R., Jones,C., Karatas,A., Larocque,K.,		
	Lamaze,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,		
	Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,		
	McManus,P., McKernan,K., McPheters,R., Meltrin,J., Meneus,L.,		
	Minova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,		
	Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,		
	Peterson,K., Phunkhang,P., Pierrre,N., Pollara,V., Raymond,C.,		
	Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,		
	Roselli,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,		
	Sever,P., Sougnez,C., Spencer,B., Strange-Thomann,N.,		
	Stojanovic,N., Strauss,N., Sudrananlan,A., Talamas,J., Tesfaye,S.,		
	Theodore,J., Travers,M., Trays,N., Triggillo,J., Vassiliev,H.,		
	Viel,R., Vo,A., Wilson,D., Wu,X., Wyman,D., Ye,W.J., Young,G.,		
	Zahoun,J., Zembek,L., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-MAR-2001) Whitehead Institute/MIT Center for Genome		
REFERENCE	Research, 320 Charles Street, Cambridge, MA 02141, USA		
AUTHORS	3 (bases 1 to 178529)		
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,		
	Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,		
	Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,		
	Chazato,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,		
	Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,		
	Fero,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D.,		
	Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,		
	Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,		
	Johnson,R., Jones,C., Kamat,A., Karatas,A., Kelis,C., Larocque,K.,		

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Dh	8519	ATCCCTGTCCCTCTCCCTGAAATTCCTACAGTCGTCCTTGGCGGGCCCTCGGAGTGGGCTT	8460
Qy	365	ValAlaValPheProValValAlaGglYThrProIlyAspSerCysSerTYrLeuCySer	384
Dh	8459	GTGGCTGTGTTCCTCCGGTGGCGGGGACCCCAAGAGAGGTGGTCTTACTGTGGCTCA	8400
Qy	385	HisThrIAsnAspArgSerIlysPheSerIleAlaAspGluAspAlaArgGlnAsnProTYr	404
Dh	8399	CACACAGCCACACAGTCCAGTTCCACATCCGGATGAAGACGACGGCAGAACCCCTTAC	8340
Qy	405	ProValIlysaIaMetGluValAlaAsnSerGlySerGluValTYrPTYrSerAsnIyPro	424
Dh	8339	CCACTGAAGCGCATGGAGGTGTCACAGGGGTCTGAGGTCTGGAGACAGCAATGGCGCG	8280
Qy	425	GlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeuGluProTYrMet	444
Dh	8279	GGCTCTCTTTCATCCGACTGTGCTCCCTCGAGANTCTGACGGCGGTGGAGGCTTACATG	8220
Qy	445	AlaProSerMetValThrSerValIlySerSerGluIlyArgGlyGluIuValVal	464
Dh	8219	GGCCCCCTCCATGGTTAGTGAAGTGTGAGCTCTGAGGGCAGAGGGAGAGGTGCTGC	8160
Qy	465	TrpCysLeuAspAspIlysaIaAsnSerLeuValMetTYrHisSerThrTYrGluIleu	484
Dh	8159	TGGTGCTCGATGACAGGCCCACTCTTGTTGATATATACACTCCACACACTACACAGCTG	8100
Qy	485	CysAlaArgTYrPheCysGlyValProSerProLeuArgAspMetPheProValArgPro	504
Dh	8099	TGTGCGCGGTACTCTCGGGGTGCCAGGCCCTCAGGAGACATGTTTCCCTGCGGGCCC	8040
Qy	505	LeuAspThrGluProProAlaAlaSerHisThrAlaAsnProIlyValProIuIlyAsp	524
Dh	8039	TTGGACACGGAAACCCCGGAGGACGACACAGCCCAACCCAAAGGTGCTTGAAGGGAC	7980
Qy	525	SerIleAlaAspValSerIleMetTYrSerGluIleuGlyThrGlnIleLeuIleHis	544
Dh	7979	TTCATCCGGACGTCGACATCATGTACAGTGAAGAGCTGGGACGAGATCTGTATCCAC	7920
Qy	545	GlnGluSerLeuThrAspTYrCysSerMetSerSerTYrSerSerSerProIlyArgGln	564
Dh	7919	CAGGATACACTGACTGACTGTGCTCATGTCTCTTACTTCTCATCCACCCACCCGCGAC	7860
Qy	565	AlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSerValProPhe	584
Dh	7859	GCTGCAGAGTCCCCCTCAAGCTCCCCAGCTCCCCAGCAAGTTCCTGATGGCTTTC	7800
Qy	585	SerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgSer	604
Dh	7799	TTCACCGACTCCGAGGACTCGACATGTCTACATACAGCCCGGCTCTCTCCACAGAGTCT	7740
Qy	605	GluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeuGlnAlaValIlyS	624
Dh	7739	GAGCATGACCTGACCCCAAGAGGAGGAGAAACCTTCAGCCAGCACTGCAAGGCCGGAAG	7680
Qy	625	IleLeuAlaValAlaArgAspLeuIleTyrValPro	635
Dh	7679	ATCCCTGCGGCTCAAGACACTCATTTGGGT-CCCAAGGTACGTTTCCGAGGTGAGGACAC	7620
Qy	635	-----	635
Dh	7620	CATCCAGGGCACGCCCACTGCTCTCTGTGGGACAGAGCAAGAGGGAAGCCCCCTGGCT	7560
Qy	635	-----	635
Dh	7560	TCCTTACAGGTCAATGGGAGGACAGGTGCTTCCAAAGAGGGTGGTTTCCCTCAT	7500
Qy	635	-----	635
Dh	7500	TGGGATGGAAGCATGCTTTAGCAAACTTAAGAGTGGCCGACAAAAGCTTAATGGCT	7440
Qy	635	-----	635
Dh	7440	GCCAGAGAGCTTAAGGTGCCACACAGAGAAAGCACAGTACATGTGACTCAGTTTCCAC	7380

```

Oy 635 ----- 635
Db 7380 TGAGAAATTCCTGCTGTTGAAGTCTGTCCCTCCCATTAATCCTTAGCAG 7321
Oy 635 ----- 635
Db 7320 AACTGCATGCTGTGCTCCTCCCGACACTTCTGCACACCGGCTTCACTCCCGGA 7261
Oy 636 ----- 639
Db 7260 GAGCAGCGCTACCGCAAAATCGCTTCCCTTCTGCTGTTTGTGCTTAGCGCGGTGA 7201
Oy 640 ASPVALIIEVALIIEGLIYEAUGLIULYASPSERGLUALAGLARGLIYARVALIIEALA 659
Db 7200 GATGTATGCTATGCTGCTCGGAGAGATTCGGGGCCGAGGGGGCCGAGCATTCGCC 7141
Oy 660 VALLEULYARLARGLIULEUTHRPROHISGLY 670
Db 7140 GTCTTAAAGCCGAGAGCTGACTCCGAGG 7108

RESULT 11
AC019254 178876 bp DNA linear HTG 06-AUG-2002
LOCUS Homo sapiens chromosome 15 clone RP11-424119 map 15, *** SEQUENCING
DEFINITION IN PROGRESS ***
AC019254
AC019254.9 GI:22123259
VERSION HTG: HTGS_PHASE1; HTGS_FULTOP; HTGS_ACTIVERIN.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 178876)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 15, clone RP11-424119
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 178876)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Balwit, J., Barne, N., Beckerly, R., Beda, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Deatellano, K., Dewar, K., Domino, M., Doyle, M., Fenesor, J.,
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Galagan, S., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howard, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., Lenocky, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPeeters, R., Meldrum, J., Menus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tjirelli, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.
DIRECT SUBMISSION
TITLE Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 178876)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barne, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chararo, B., Choepel, L., Collymore, A.,
Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Galagan, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamet, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Risse, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,

```

```

TITLE
JOURNAL
COMMENT
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
DIRECT SUBMISSION
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:21699484.
All repeats were identified using RepeatMasker.
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: I2454
Center clone name: 424_1.19
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 58973: contig of 58973 bp in length
* 58974 59073: gap of 100 bp
* 59074 91153: contig of 32080 bp in length
* 91154 91253: gap of 100 bp
* 91254 178876: contig of 87623 bp in length.
Location/Qualifiers
1. 178876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-424119"
/clone_11b="RP11-424119 Human Male BAC"
BASE COUNT 48851 a 44853 c 42570 g 42356 t 246 others
ORIGIN
Alignment Scores:
Pred. No.: 2,15e-102 Length: 178876
Score: 1560.50 Matches: 319
Percent Similarity: 68.15% Conservative: 2
Best Local Similarity: 67.73% Mismatches: 4
Query Match: 40.83% Indels: 147
DB: 2 Gaps: 2
US-09-836-392-21 (1-728) x AC019254 (1-178876)
Oy 346 Valproval-----Ilelysyasnseryrlyuvalleuallaglyleuallaapglyeu 364
Db 14759 ATCCCTGTCTCTCTCTTGAATTCCTACCGGCTTAGCGGCTCCCATGAGGCTT 14818
Oy 365 VALALAVAlpheProvalValalArglyThrPolysAspSerCysseryrleucySer 384
Db 14819 GTCGCTGTTCCTCCCTGTCGGGGGACCCCAAGACAGCTGCTCCTACTGTCCTA 14878
Oy 385 HASTHAlaAsnArgSerlyrYspHeserIlealaAspGluAspAlaArgGlnAsnProtyr 404
Db 14879 CACACAGCCACAGGTCACAAAGTTTCAGCATCGGATGAAGACGACGAGAAACCCCTAC 14938
Oy 405 PROVALYAlaMetGluValValAsnSerGlyserGluValITrptysersnGlyPro 424
Db 14939 CCAGTGAAGCCATGAGGTGTGCAACACCGGCTGTGAGTGTGTACAGCAATGGCGCG 14998
Oy 425 GYLLEULYVALIIEAspCyAlaSerLeuGluIIECyArGArgLeuGluProTyrmec 444
Db 14999 GGCTTCCTGTGATGACAGCTGCTCCCTCGAGATCTGACAGCGGCTGAGCCCTACATG 15058

```


COMMENT

University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@elms.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'-6' end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

location/Qualifiers

1..2047
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="LNG07978"
/tissue-type="human lung"
/clone_id="LNG"
/note="cloning vector pME18FL3"
21..782
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB1547.1"
/db_xref="GI:10439701"
/translation="MYSTTYOLCARYCGVPSLRDMPYRPLDTPPPASHTANPK
VPKDSIAIVSIVSEIETGQIILHOSLIDYCSMSYSSPPROAARPSLSPSA
SSSVPTSTDCSDMDHTPRASDRSEHDLTPDGETFSCHLOAKITLARDLITWP
RRGSDVIVIGLEKSGAQRGRVAVLAKRELTTPRGVAVDAVAVAKDVVCTFENENE
WCLAVMNGMGRREDIDFYQSYEEELGRLEACTRRKR"

CDS

BASE COUNT 481 a 574 c 536 g 456 t
ORIGIN

Alignment Scores:

Pred. No.: 9.11e-90 Length: 2047
Score: 1352.00 Matches: 258
Percent Similarity: 99.61% Conservative: 0
Best Local Similarity: 99.61% Mismatches: 1
Query Match: 35.37% Indels: 0
DB: 9 Gaps: 0

US-09-836-392-21 (1-728) x AK026772 (1-2047)

QY 470 LysAlaAsnSerLeuValMetThrHisSerThrThyArgLysLeuGlyAlaArgTyrPhe 489
DB 3 AAGGCCAACTCTGTGTGTACACACACACACACACACACACACACACACACACACAC 62
QY 490 CysGlyValProSerProLeuArgAspMetPheProValArgProLeuAspThrGluPro 509
DB 63 TCGGGGTCGCCAGCCCTCGGACACATGTTCCCGTGGCCCTTGACACAGGACCC 122
QY 510 ProAlaAlaSerHisThrAlaAsnProLysValProGluGlyAspSerIleAlaAspVal 529
DB 123 CCGGACAGCAGCCACACAGCCCAACCAAGGCTGAGGGGAGCATTCATCCGACGATG 182
QY 530 SerIleMetTyrSerGluGluLeuGlyThrGlnIleLeuIleHisGlnGluSerLeuThr 549
DB 183 AGCATCATGATACAGTACGAGAGCTGGGACGACGATCTGATCCACAGGAATCATCTACT 242
QY 550 AspTyrCysSerMetSerSerTyrSerSerSerProProArgGlnAlaAlaArgSerPro 569
DB 243 GACTACTGCTCATCT 302
QY 570 SerSerLeuProSerSerProLysSerSerSerValProPheSerThrAspCysGlu 589
DB 303 TCAACCTCCACAGCTCCACACAGATTCTTCCAGTGCCTTCTCCACGACGACGAG 362
QY 590 AspSerAspMetLeuHisThrProGlyAlaAlaSerAspArgSerGluHisAspLeuThr 609
DB 363 GACTCAGACATCTACATACGCCCGTGTCTCCGACAGCTGAGCATGACCTGACCC 422
QY 610 PrometAspGlyGluThrPheSerGlnHisGlnAlaValIleLeuAlaValAlaGly 629
DB 423 CCCATGAGAGGGGAGACCTTACGACGACCTGACAGGCCGCTGAAGATCTCTCCCGTCAGA 482

QY 630 AspleuIleTyrValProArgArgGlyGlyAspValIleValIleGlyLeuGlyLysAsp 649
DB 483 GACTCATTTGGTCCCGCCAGCGGGTGGAGATGTATTCGTCATTTGGCTGGAGAGAT 542
QY 650 SerGluAlaGlnAAGGlyArgValIleAlaValLeuLysAlaArgGluLeuThrProHis 669
DB 543 TCTGGCGCCAGCGGGGCGGAGTATTCGCTTTAAAGCCGAGACCTGACCTCCGAT 602
QY 670 GlyValLeuValAspAlaAlaValAlaLysAspThrValValCysThrPheGluAsn 689
DB 603 GGGGTGCTGGTGGATGCTGCGGTGGTGGCAAGGACACTGTGTGTGCACCTTGAAAT 662
QY 690 GluAsnThrGluTTPCysLeuAlaValIlePheArgGlyTTPGlyAlaArgGluPheAspIle 709
DB 663 GAAACACAGAGTGTGCTGCTGCGCTGCGAGGGCTGGGCGCCAGGAGATTGACATT 722
QY 710 PheTyrGlnSerTyrGluGluLeuGlyArgLeuGluAlaCysThrArgLysArgArg 728
DB 723 TTCTACAGTCTCTACGAGAGACTGGGCGGCTGGAGGCTTGACATCCCAAGAGAGG 779
RESULT 14
AC127595 288428 bp DNA linear HTG 11-AUG-2002
LOCUS AC127595
DEFINITION Mus musculus chromosome UNK clone RP23-49118, WORKING DRAFT
SEQUENCE 12 unordered pieces.
AC127595 AC127595.2 GI:22203893
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 288428)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE The sequence of Mus musculus clone Unpublished
2 (bases 1 to 288428)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
2 Submitted (17-JUL-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 288428)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
Submitted (11-AUG-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
On Aug 11, 2002 this sequence version replaced gi:21887033.
COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BA0049118
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 28337 bases at least Q40
Consensus quality: 284218 bases at least Q30
Consensus quality: 284790 bases at least Q20
----- NOTE: This is a "working draft" sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence.

Db 191250 GGGGAGACGTTGACGACGACCTGACAGGCTGTGAGAGTCTGTCTCAAGACATCAT 191309
 QY 633 TTPVALPRARG----- 637
 Db 191310 TGGGTTCTAGGTA-AGCTTTGTGTGTGAGAGTATCATTTCTGAGCCCTTCTTGGG 191368
 QY 637 ----- 637
 Db 191369 GCTGGGGTAGTGAACCCGTGCTCTCTTAAGATTCCAGSTCCATCTGGCCGTTCT 191428
 QY 637 ----- 637
 Db 191429 GCCGAGACGCTTGCCATTCGTTTAGAATTGAGAGTATCTGATAGACAGATCC 191488
 QY 638 ----- 638
 Db 191489 TGAGCAGCTCTTCCCTGCTTGTATTTGGTCCCTTAGACATGAGAGTATCATCTGTC 191548
 QY 644 TLeGLLeuGLuLysAspSerGLuAlaArgGLuArgValLLeuValLeuLysAla 663
 Db 191549 ATCGGCTGGAGAGATTCAGGTGCCCGAGGGGCAAGTCTCGCTTTTAAAGCC 191608
 QY 664 ArgGLuLeuThrProHisGLy----- 670
 Db 191609 CGAGAGCTCAATCGGACGGGTAAAGTGAACCTTCTATATGGCTGGGGGAGGGGT 191668
 QY 670 ----- 670
 Db 191669 GTGGGAGACCTTAGACATCACTACATGCCACTAGACAGACAGAGCCCTGCTACCCC 191728
 QY 670 ----- 670
 Db 191729 CTGGGCTCCACAGATTAAGCTCTGGATCCACAGTGTGGCCACAGAGTCAAAAGAGC 191788
 QY 670 ----- 670
 Db 191789 AAGTAGGGCTCTCAAGGACGCTCGAGAGTTCAGCATGTGTGCTGAATAGAGATCAAC 191848
 QY 670 ----- 670
 Db 191849 CACTGGCAAGATATCTATTCTGGGCTGTGTGACTGTGATCAAGTACACTGCC 191908
 QY 670 ----- 670
 Db 191909 AAGGTGTATCACAGGGCCAGATTCAGCTGGATTGAAAGAAAGTCTCTCCAAATTGGT 191968
 QY 670 ----- 670
 Db 191969 TCAGTGTCTCTCTCCAAACAGAGAGTCCGCCACCCCAATGTTAGGGAGACCCACA 192028
 QY 670 ----- 670
 Db 192029 GCCCCATGAAGCAGATAGGGCCAGATGAGATGGCCCACTCTGACGCTCTCTC 192088
 QY 670 ----- 670
 Db 192089 TGAGCTGAGAGTGTGACTGTATCTAATCTGAGTAGAGCTAGAGACATGCCAG 192148
 QY 670 ----- 670
 Db 192149 TCGCTGTGTGAGAGACAAGTCCAGGGCAGACAGCTTTGTGGCCAGGAGATGGAGCT 192208
 QY 670 ----- 670
 Db 192209 GGCATTAAAGCAGAGTCTTGCGAGTGTATCCCTTACCCGAAAGTCCCATGAAGT 192268
 QY 670 ----- 670
 Db 192269 CCCTAAAGCCAAAGTGTGAATTTCCATGCTAGATTCATCTGCGAGCATGGCT 192328
 QY 670 ----- 670
 Db 192329 GTGACTCATCTTGAGCAAAAGTCTTGATTCATCTTTAATATCATGCGACAGTCCCATTC 192388

QY 670 ----- 670
 Db 192389 TGCCCAAGACAGAGACATGATGACAAACAGCCAGTCACTCATGCCATCCAGTGGCTT 192448
 QY 670 ----- 670
 Db 192449 GCCTGTAGTAGTCTGGGATCTCTCTGTGAATAAGATCAGAGCTTTGCTCGAGCC 192508
 QY 670 ----- 670
 Db 192509 TGCGTACCTGGCCAAAGAGCTTGTGTCCAAGCTGGGCTTTGGGAAACAATGGTTC 192568
 QY 670 ----- 670
 Db 192569 CTCAGAACAGAGTTGGTTGACCTTGGCAATACCTACTGTACTATCTGATCTCTTTCT 192628
 QY 670 ----- 670
 Db 192629 AGTCACCCCCCAACCCCAACCCCTGAGACAGSTTTTATCAAGAGTGGGCCACGTTGTGTA 192688
 QY 670 ----- 670
 Db 192689 AGCAGAAAGACCCAGTGTGGCCAGTAAAGAAAGAAAGAGACAGACTCTGTATAC 192748
 QY 670 ----- 670
 Db 192749 CTTGGTAGAGAGAGCTTCCTCATGTGAGTGGGCTCAGTCACTGGGCTTAGAGAGAG 192808
 QY 670 ----- 670
 Db 192809 ATACAAACAGACCCAGATGAGTGGCCAGCCCTCGCAGAGACACATCAGAGACAGAGA 192868
 QY 670 ----- 670
 Db 192869 ATGCTGAGAGAAACTTGCTATGATGACAGAGCCATGACAAACAGCTCACAGCTCA 192928
 QY 670 ----- 670
 Db 192929 GAAACAAAGTAAGAGCCGCTGCTGAGATAGATCTTCCTGCTTAACATATCCCCC 192988
 QY 670 ----- 670
 Db 192989 CTGTCTCTGAGTCTGTATCTTCCAGAGCGTTAGTCTACAGCAATGCTTGCAAAAGTA 193048
 QY 670 ----- 670
 Db 193049 AAGGGTACCCCTTGGCCTTCAGGTATCATTTAAGTCCCATCCATAGGCCGGAATTACC 193108
 QY 670 ----- 670
 Db 193109 AGGTAGATCAGGACTTCCAAAGTTCAATCTAGAGCCAGCTGAGCTGTTATGACACCTT 193168
 QY 670 ----- 670
 Db 193169 CTTAGTACTCCGCCCTTCAAGATATCCATGAAGAGGAGGCCCACTGACCTAG 193228
 QY 671 ----- 671
 Db 193229 TGAGGACTCTGTGGGTGATGATAGATCCCATCTCTGCAAGGCTTGTGGAGATGCC 193288
 QY 676 ALAValAlaLysAspThrValValCysThrPheGLuAsnGLuAsnThrGLuTPCys 695
 Db 193289 GCAGTGTGTGCAAAAGGACACCGTAGTGTGAGCTTTCACAAATGAACACAGAGTGTGC 193348
 QY 696 LeuAlaValTrpArgGLuTrpGLuValArgGLuPheAspLLeuPheThrGLuInsThrGLu 715
 Db 193349 CTGGCTGTCTGGAGGGGCTGGGCGCCAGGAGTTGATCTTCTTACAGCTCTATGAG 193408
 QY 716 GLuLeuGLuArgLeuGLuAlaCysThrArgLysArgArg 728
 Db 193409 GAGCTGGGCGGCTGGAGGCTTGTGACTCGCAAGAGGAG 193447

RESULT 15
ARI75612 ARI75612 526 bp DNA linear PAT 17-DEC-2001
LOCUS ARI75612
DEFINITION Sequence 10 from patent US 6309849.
ACCESSION ARI75612
VERSION ARI75612.1 GI:17916911
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 526)
AUTHORS Robison, K.E.
TITLE Nucleic acid molecules encoding human kinase and phosphatase
JOURNAL Patent: US 6309849-A 10-30-OCT-2001;
FEATURES Location/Qualifiers
source 1..526
BASE COUNT 115 a 137 c 160 g 111 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 3,17e-54 Length: 526
Score: 856.00 Matches: 169
Percent Similarity: 97.69% Conservative: 0
Best Local Similarity: 97.69% Mismatches: 1
Query Match: 22.40% Indels: 3
DB: Gaps: 0
US-09-836-392-21 (1-728) x ARI75612 (1-526)
QY 99 AspleuysSerAspAsnIleuValTrpSerLeuAspValLysGluHisIleAsnIle 118
DB 12 GACCTGAAGTCGGACAACTCTGGTGGTCCCTTGACGTCACAGACACATCAACATC 71
QY 119 LysLeuSerAspTyrGlyIleSerArgInsPheHisGluGlyAlaLeuGlyValGlu 138
DB 72 AAGCATATGACTACGAGCATTCGAGGCATTCATGAGGCGCCCTAAGCGCTGGAG 131
QY 139 GlyThrProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysVal 158
DB 132 GGCACCTCTGGTACACAGCCCAAGATCAGGCTCGCATGTATGATGAGAAAGTA 191
QY 159 AspMetPheSerTyrGlyMetValLeuTyrGluLeuSerGlyGlnArgProAlaLeu 178
DB 192 GATATGTTCTCTATGAGATGGTCTCTACGAGTTGCTCTCAGACAGCGCCCTGCACCTG 251
QY 179 GlyHisHisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGly 198
DB 252 GGCACACACAGCTCCATTCGCAAGAGCTGCCAAGGCGCATCCGCCGCTTCTGGGG 311
QY 199 GlnProGluGluValGlnPheArgArgLeuGlnAlaLeuMetGluCysTrpAspThr 218
DB 312 CAGCCGGAGAGAGTGCATCTCGGCGACTGCAGCGCTCATGATGAGTGGTGGACACT 371
QY 219 LysProGluLysArgProLeuAlaLeuSerValVal-SerGlnMetLysAspProThrPh 238
DB 372 AAGCCAGAGAAAGCAGCACTGGCTGCTGGTGGTGAAGCCAGATGAAGGAGCCGACTTT 431
QY 238 eAlaThrPheMetTyrGluLeuCysGlyLysGlnThrAlaPhePhe-SerSerGlnG 258
DB 432 TGCACCTTCATGATGAACTGTCTGTGGGAGAGCAGACAGCTTCTTCTCATNCCAG 491
QY 258 LysGlnGluTyrThrVal-ValPheTrpAspGly 268
DB 492 GCCAGAGTACACTGTGGGTGTTTGGGATGGA 524

Search completed: April 15, 2003, 08:37:18
Job time : 3982 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 15, 2003, 07:15:13 ; Search time 283 seconds

(without alignments)
5793.126 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822

Sequence: 1 MRLHRLATDAMKNSEFROE.....IFYSYELGLRLEACTRRKR 728

Scoring table: BLOSUM62

Xgapop 10.0, Ygapext 0.5

Xgapop 10.0, Ygapext 0.5

Ygapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing: first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/cgn2_1/USPTO.spool/US0936392/runat_08042003_090539_23389/app_query.fasta.1.903

-DB=Geneseq.101002 -QEXT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0

-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosm62 -TRANS=human40.cdi

-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFILE=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USPR=US0936392.cgn.1.1.200 &runat_08042003_090539_23389 -NCPU=6 -ICPU=3

-NO.XLPRX -NO.MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120

-MARK.TIMEOUT=30 -THEADS=1 -XGAPOP=10 -YGAPEXT=0.5 -Fgapop=6 -Fgapext=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

1: Geneseq.101002:*

2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*

24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3822	100.0	3496	AA03995	Human protein tyro
2	3822	100.0	3579	AA033718	Human secreted pro
3	3815	99.8	3564	AA033695	Human secreted pro
4	3728	97.5	2652	AA072238	DNA encoding novel
5	3579.5	93.7	6045	AA077994	Nucleotide sequenc
6	3579.5	93.7	7093	AA038851	Human kinase (PKIN
7	1360.5	40.8	18286	AAK70889	Human kinase (PKIN
8	1560.5	40.8	18286	AAK84104	Human kinase (PKIN
9	1560.5	40.8	18286	AAK70892	Human kinase (PKIN
10	1560.5	40.8	18286	AAK84107	Human kinase (PKIN
11	856	22.4	526	AA017057	Human protein kin
12	404	10.6	9947	AB09444	Drosophila melanog
13	364.5	9.5	6927	AB09445	Drosophila melanog
14	330.5	8.6	2693	AA026452	Human kinase PKIN-
15	330.5	8.6	2730	AA06754	Poly nucleotide seq
16	315	8.2	1631	AA039537	Arabidopsis thalia
17	314	8.2	1224	AA043254	Arabidopsis thalia
18	307.5	8.0	3454	AB070018	Pancras cancer re
19	288.5	7.5	3072	AB04365	Drosophila melanog
20	288	7.5	2283	AA048526	Arabidopsis thalia
21	276.5	7.2	556	AA099418	Human protein enco
22	273	7.1	3141	AA018824	Human kinase (PKIN
23	270.5	7.1	3538	AA034309	Human PKIN-12 CDNA
24	269.5	7.1	3111	AB06357	Novel human protei
25	269.5	7.1	3518	AB06358	Novel human protei
26	268	7.0	3066	AB066165	Novel human gene.
27	263	6.9	2157	AA046913	CDNA encoding huma
28	260.5	6.8	3558	AB03874	Human CDNA differe
29	257.5	6.8	3558	AA036139	Human mtogen acti
30	257.5	6.7	3389	AA01031	Human leucine zipp
31	257.5	6.7	3389	AB09349	Human leucine zipp
32	256.5	6.7	3365	AB084203	Human CDNA differe
33	253.5	6.6	1514	AA047638	Arabidopsis thalia
34	251.5	6.6	3048	AA046914	CDNA encoding huma
35	249	6.5	3281	AA090484	Arabidopsis thalia
36	249	6.5	4686	AB011318	Drosophila melanog
37	246.5	6.4	4419	AB044653	Drosophila melanog
38	246	6.4	4380	AB015973	Lung cancer relate
39	243.5	6.4	2863	AB066321	PTK gene LPTK-2.
40	243.5	6.4	7607	AA049754	Protein tyrosine-k
41	243.5	6.4	7607	AA03097	CDNA encoding huma
42	243	6.4	1365	AA029727	DNA encoding a hum
43	243	6.4	1706	AA045674	CDNA encoding huma
44	243	6.4	2120	AA029726	Human protein enco
45	243	6.4	2191	AA099263	

ALIGNMENTS

RESULT 1
AA03995
AA03995 standard; CDNA; 3496 BP.

02-JUL-2001 (first entry)

Human protein tyrosine kinase receptor (PTK) CDNA from clone HDPS68.
Human; protein tyrosine kinase receptor; PTK; antibacterial; antiviral;
antifungal; antiparasitic; sarcoidosis; inflammation; asthma; arthritis;
cardiovascular disorder; severe combined immunodeficiency; SCID; vaccine;
hyperproliferative disorder; atherosclerosis; brain disorder; leukemia;
multiple sclerosis; Alzheimer's disease; vascular dementia; thrombosis;
neurodegenerative disorder; gastrointestinal disorder; cancer; lymphoma;
blood coagulation disorder; trauma; cerebrovascular disorder; tendonitis;
gene therapy; nephrotic syndrome; glomerulonephritis; allergy; neoplasm;

KW musculo-skeletal disorder; Parkinson's disease; autoimmune disorder;
 KW behavioural disorder; renal disorder; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 22..2208
 FT /tag="a
 FT /product="Human PTK protein from clone HDP5B68"
 PN WO200129564-A1.
 PD 26-APR-2001.
 PF 12-OCT-2000; 2000MO-US28066.
 PR 15-OCT-1999; 99US-0159542.
 PR 17-NOV-1999; 99US-0165914.
 PR 14-MAR-2000; 2000US-0189027.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Shi Y, Young PE, Ni J;
 DR WPI: 2001-290976/30.
 DR P-PSDB: AAE00669.
 XX
 PT New isolated nucleic acid molecule for producing human protein tyrosine
 PT kinase receptor
 PS Claim 1; Page 265-266; 288bp; English.
 CC The present cDNA sequence encodes human protein tyrosine kinase receptor
 CC (PTK) from clone HDP5B68.
 CC The invention relates to human protein tyrosine kinase (PTK) receptors
 CC and their corresponding cDNA molecules. PTK receptors are useful in
 CC providing immunological probes for differential identification of the
 CC tissues or cell types present in a biological sample. PTK is used in
 CC methods for the diagnosis, prevention and treatment of various disorders
 CC related to PTK such as immune system disorders (severe combined
 CC immunodeficiency (SCID), inflammation; hyperproliferative disorders
 CC (neoplasia, sarcoidosis); cardiovascular disorders (arrhythmia,
 CC atherosclerosis); central nervous system disorders (multiple sclerosis),
 CC neurodegenerative disorders (Alzheimer's disease, Parkinson's disease);
 CC blood coagulation disorders (thrombocytopenia); autoimmune disorders
 CC (biliary cirrhosis, Crohn's disease); respiratory disorders (asthma,
 CC allergy); gastrointestinal disorders (inflammatory bowel disease);
 CC cerebrovascular disorders (thrombosis, vascular dementia); brain
 CC disorders (phenylketonuria); cancers (such as ovarian, lung, bladder,
 CC liver, breast and lymphomas); behavioural disorders (Tourette's
 CC syndrome); musculo-skeletal disorders (arthritis, trauma, tendonitis);
 CC renal disorders (nephrotic syndrome, glomerulonephritis); metastases of
 CC malignancies and related disorders (leukemia, multiple myeloma); and
 CC infections caused by bacteria, viruses, fungi and parasites. PTK is also
 CC useful for screening therapeutic compounds. PTK is used as an antigen in
 CC a vaccine to raise an immune response against infectious disease. PTK
 CC nucleic acids are useful in gene therapy.
 XX
 SQ Sequence 3496 BP; 812 A; 998 C; 949 G; 737 T; 0 other;
 Alignment Scores:
 Pred. No.: 3,926-295 Length: 3496
 Score: 3822.00 Matches: 728
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-836-392-21 (1-728) x RAD03995 (1-3496)
 QY 1 MettLeuArgHisLeuAlaThrAspAlaMetLysAsnPhsSerGluPheArgGlnGlu 20
 DB 22 ATGTGAGGACCTGCGGGCCACCGATGCCATGAAGAACTCTCCGAGTTCGGCGAGAG 81

QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
 DB 82 GCCAGCATCTGACGCGCTGCACACCCCTGCATCTGCGCTCTACGGCATCAGCATC 141
 QY 41 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 DB 142 CACCGGCTGCTGCGCCCTGGAGCTCGCGCGCTGACAGCATCAACCGCTGCTGCTC 201
 QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnIleAla 80
 DB 202 GAGAACGCCAGATCTTCTTATATACCCCTGGGACACATCTCACCAAAAAATAGCC 261
 QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysAsnIleIlePheCysAspLeu 100
 DB 262 TACCAATGCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
 QY 101 LysSerAspAsnIleLeuValTyrPheLeuAspValLysGlnHisIleAsnIleLysLeu 120
 DB 322 AAGTGGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
 QY 121 SerAspTyrGlyIleSerArgGlnSerPheHisGlyAlaLeuGlyValGluGlyThr 140
 DB 382 TCTGACTACGGGATTTGAGGACATCTCATGAGGGCGCCCTAGGCGCTGAGGCGACT 441
 QY 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet 160
 DB 442 CCGGCTACACGCGCCCAAGATCAGGCTCCCATTTGATGATGAGAGGTAGATTAAG 501
 QY 161 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHis 180
 DB 502 TCTCCTATGAGAAATGCTGCTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
 QY 181 HisGlnLeuGlnIleAlaLysLeuSerLysGlyIleArgProValLeuGlnPro 200
 DB 562 CACCGAGCTCAATTCACAAAGAGCTGCTCAAGAGGCTCCCGGCTGCTGCTGCTGCTGCT 621
 QY 201 GluGlnValGlnPheArgArgLeuGlnAlaLeuMetGlnCysTyrPheArgThrAspPro 220
 DB 622 GAGGAGTCAAGTTCGCGGAGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
 QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
 DB 682 GAGAGCGACCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
 QY 241 PheMetTyrGlnLeuCysGlyGlnThrAlaPhePheSerSerGlnGlnGlu 260
 DB 742 TTCATGTATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
 QY 261 TyrThrValValPheThrAspGlyLysGlnGlnArgAsnThrThrValValAsnThr 280
 DB 802 TACACCGTGTGTTTGGATGAAAGAGAGTCCAGGACTACCGGTGTGAACACA 861
 QY 281 GluLysGlyLeuMetGlnValAlaGlnArgMetCysCysProGlyMetLysValSerCysGln 300
 DB 862 GAGAGGCGCTCATGAGAGTGCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
 QY 301 LeuGlnValGlnArgSerLeuThrAlaThrGlnAspGlnLysIleTyrIleTyrThr 320
 DB 922 CTCACAGTCCAGAGATCCCTGTGACAGCCAGGACCCAGAAATATCATCATCAACC 981
 QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValVal 340
 DB 982 CTCACAGGAGTATGCTGCTTAAACACCCCAAGAGGCTTGGATACCTCACCTGCTGCTGCT 1041
 QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuIleGlyLeu 360
 DB 1042 ACCTGCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
 QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 DB 1102 GCCGATGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161

QY 381 TTYRLeuCySerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
 DB 1162 TACCTGTCTCACACACACAGCAGCAGTCAAGTTCAGCATCGGATCAAGACGACG 1221
 QY 401 GlnAsnProTyrProValLysAlaMetGluValAlaAsnSerGlySerGluValTyrTyr 420
 DB 1222 CAGAACCCCTACCCAGTGAAGCCATGAGAGTGTGTCAACAGCGGCTCTAGGCTGTGAC 1281
 QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgLeu 440
 DB 1282 ACCAATGTGGCCGGCTCTCTTGCATCGACTGTGCTCCCTCGAGATCTGACGGCGGTG 1341
 QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGluGlyArgGly 460
 DB 1342 GAGCCCTACATGCCCCCTCCATAGTTCAGTGTGTGTGACGTCTGAGGGCAGAGGG 1401
 QY 461 GluGluValValTyrPCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 DB 1402 GAGGAGGT 1461
 QY 481 ThrTyrGluLeuCyAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500
 DB 1462 ACTTACACACTGT 1521
 QY 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
 DB 1522 CCGGTGCGCCCTGTGACACAGGAAACCCGCGACGACGACGACGACGACGACGACGACG 1581
 QY 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluGluLeuGluGlyThrGln 540
 DB 1582 CCGGAGGGGAGACCCATCCGCGAGCGAGCATGATGATGATGATGATGATGATGATGATG 1641
 QY 541 IleLeuIleHisGlnIleuSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSer 560
 DB 1642 ATCTGATCTCACAGGAAATCCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGCT 1701
 QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
 DB 1702 CCACCCGCGACGCTGCGAGGTCCCTCCATAGCCTCCCTCCATAGCCTCCCTCCATAGC 1761
 QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
 DB 1762 AGTGTGCTTTCCTCCACGACGCTGCGAGGACTGACATGATGATGATGATGATGATGAT 1821
 QY 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
 DB 1822 TCCGACAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1881
 QY 621 GlnAlaValLysIleLeuAlaValArgAspLeuIleTyrValProArgArgGlyAsp 640
 DB 1882 CAGGCGGTGAAGATCTCGCTGACGAGACCTATTGGGTCCCGCAGGCGGCGGTGGAGAT 1941
 QY 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 DB 1942 GTTATGTGATTTGGCGCGGAGGAGATTCGAGGCCAGCGGGGCGGAGCATTCGCGCTG 2001
 QY 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValValAlaLys 680
 DB 2002 TTTAAAGCCCGAGAGGTGATCCGCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2061
 QY 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrCysLeuAlaValTyrArg 700
 DB 2062 GACACTGTTTGTGACCTTTGAAATGAAACACAGAGTGGTGGCTGGCGCTGTGAGG 2121
 QY 701 GlyTyrPolValaArgGluPheAspIlePheTyrGlnSerTyrGluGluGlyArgLeu 720
 DB 2122 GGGTGGGGGCGCAGAGGAGTTCGACATTTTCTACAGTCCACAGAGGAGCTGGCGGCTG 2181
 QY 721 GluAlaCysThrArgLysArgArg 728
 DB 2182 GAGGCTTGCTGCTGCAAGAGG 2205
 RESULT 2

AAD33718
 ID AAD33718 standard. CDNA: 3579 BP.
 XX
 AC AAD33718;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human secreted protein-encoding gene 4 CDNA clone HDMW90. SEQ ID NO:37.
 XX
 KW Human; secreted protein; immune disorder; antiallergic; antirheumatic;
 KW rheumatoid arthritis; breast neoplasia; breast cancer; antiallergic;
 KW neurological disease; Alzheimer's disease; Parkinson's disease; trauma;
 KW Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania;
 KW antinflammatory; ophthalmological; dermatological; immunostimulatory;
 KW immunomodulatory; immunosuppressive; antibacterial; antiparasitic;
 KW gene therapy; autoimmune disease; Huntington's disease; meningitis;
 KW demyelinating disease; peripheral neuropathy; congenital malformation;
 KW spinal cord injury; peripheral neuropathy; ischemia; perception;
 KW multiple sclerosis; infection; haemorrhage; schizophrenia; dementia;
 KW depression; panic disorder; learning disability; AIDS; feeding disorder;
 KW hyperproliferative disorder; sleep pattern; cardiovascular disorder;
 KW reproductive disorder; digestive system disorder; behavioural disorder;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 134..2254
 FT /tag= a
 FT /product= "Human secreted protein"
 FT sig_peptide 134..205
 FT /tag= b
 FT mat_peptide 206..2251
 FT /tag= c
 FT /product= "Mature secreted protein"
 XX
 PN MO200216390-A1.
 XX
 PD 28-FEB-2002.
 XX
 PF 17-JAN-2001; 2001WO-US01435.
 XX
 PR 18-AUG-2000; 2000US-226282P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
 PI Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fischella M;
 PI Ni J;
 XX
 DR WPI: 2002-304113/34.
 DR P-PSDB: AAE21217.
 XX
 PT An isolated nucleic acid molecule (1) comprising a polynucleotide which
 PT encodes a polypeptide useful in the diagnosis and treatment of
 PT disorders e.g. immune disorders -
 XX
 PS Claim 1: Page 457-458; 534p; English.
 XX
 CC AAD33692-AAD33736 represent CDNA corresponding to 21 human secreted
 CC protein genes, and AAE21191-AAE21235 represent the proteins they encode.
 CC AAE21236-AAE21280 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 21 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC immune or autoimmune diseases e.g. AIDS (acquired immune deficiency
 CC syndrome), asthma, anaemia and rheumatoid arthritis, breast neoplasia
 CC and breast cancer, neurological diseases e.g. Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, Tourette syndrome,

QY 621 GlnAlaValIleLeuAlaValArgAspLeuIleTrpValProArgArgGlyGlyAsp 640
 DB 1928 CAGGCCGGAGATCTCCGCCGCGAGACCTCTTTGGTCCCCAGCGCGTGAGAT 1987
 QY 641 ValIleValIleGlyLeuGlyAspSerGlnIleGlnArgGlyArgValIleAlaVal 660
 DB 1988 GTTATCGCATGGCGCTGGAGAGATCTGAAGCCGAGCGGGCGGCGATTCGCCGTC 2047
 QY 661 LeuIleValArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaVal 680
 DB 2048 TTTAAAGCCGAGAGCTACCTCCGATGGGTGCTGTGATGCTGCGTGTGCAAG 2107
 QY 681 AspThrValValCysThrPheGlnAsnGlnAsnThrIleTrpCysLeuAlaValTrpArg 700
 DB 2108 GACACTGTGTGTGTCACCTTTGAAAATATAACACAGAGTGCTGCTGCGCTGAGAG 2167
 QY 701 GlyTrpGlyAlaArgGluPheAspIlePheThrGlnSerTrpGlnGluLeuGlyArgLeu 720
 DB 2168 GGCTGGGGCGCCAGGAGTTCGACATTTCTACACAGTCTACGAGAGAGCTGGCGCGCTG 2227
 QY 721 GlnAlaCysThrArgLeuArgArg 728
 DB 2228 GAGGCTTGCACTCCGCAAGAGAAG 2251

RESULT 3
 AAD33695
 ID AAD33695 standard; CDNA: 3564 BP.
 AC AAD33695;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human secreted protein-encoding gene 4 CDNA clone HDDMW90, SEQ ID NO:14.
 XX
 KW Human; secreted protein; immune disorder; anti-allergic; anti-rheumatic;
 KW Rheumatoid arthritis; breast neoplasia; breast cancer; anti-arthritic;
 KW neurological disease; Alzheimer's disease; Parkinson's disease; trauma;
 KW Tourette syndrome; encephalitis; cytostatic; hemostatic; anaemia; malaria;
 KW anti-inflammation; ophthalmological; dermatological; immunostimulatory;
 KW Immunomodulatory; immunosuppressive; antibacterial; antipsoriatic;
 KW gene therapy; autoimmune disease; Huntington's disease; meningitis;
 KW demyelinating disease; peripheral neuropathy; congenital malformation;
 KW spinal cord injury; peripheral neuropathy; ischaemia; perception;
 KW multiple sclerosis; infection; haemorrhage; schizophrenia; dementia;
 KW depression; panic disorder; learning disability; AIDS; feeding disorder;
 KW hyperproliferative disorder; sleep pattern; cardiovascular disorder;
 KW reproductive disorder; digestive system disorder; behavioural disorder;
 KW gene; ss.
 KW
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 140..2260
 FT /tag= a
 FT /product= "Human secreted protein"
 FT sig_peptide 140..211
 FT /tag= b
 FT mat_peptide 212..2257
 FT /tag= c
 FT /product= "Human mature secreted protein"
 XX
 PN WO200216390-A1.
 XX
 PD 28-FEB-2002.
 XX
 PE 17-JAN-2001, 2001WO-US01435.
 XX
 PR 18-AUG-2000, 2000US-226282P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;

PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
 PI N1 J;
 XX
 DR WPI: 2002-304113/34.
 DR P-PSDB: AAE21194.
 XX
 PT An isolated nucleic acid molecule (1) comprising a polynucleotide which
 PT encodes a polypeptide useful in the diagnosis and treatment of
 PT disorders e.g. immune disorders -
 XX
 PS Claim 1; Page 442-443; 534pp; English.

CC AAD33692-AAD33736 represent CDNA's corresponding to 21 human secreted
 CC protein genes, and AAE21191-AAE21235 represent the proteins they encode.
 CC AAE21236-AAE21280 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 21 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC immune or autoimmune diseases e.g. AIDS (acquired immune deficiency
 CC syndrome), asthma, anaemia and rheumatoid arthritis, breast neoplasia
 CC and breast cancer, neurological diseases e.g. Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, Tourette syndrome,
 CC meningitis, demyelinating disease, peripheral neuropathies, neoplasia,
 CC trauma, congenital malformations, spinal cord injuries, toxic
 CC neuropathies induced by neurotoxins, peripheral neuropathies, multiple
 CC sclerosis, ischaemia and infarction, haemorrhages, schizophrenia, mania,
 CC dementia, depression, panic disorder, learning disabilities, AIDS,
 CC altered behaviours e.g. disorders in feeding, sleep patterns, balance
 CC and perception, encephalitis, disorders in cardiovascular, neural/
 CC sensory, reproductive and digestive systems, behavioural disorders and
 CC hyperproliferative disorder. The present sequence represents a human
 CC secreted protein-encoding CDNA of the invention.
 CC
 XX
 SQ Sequence 3564 BP; 844 A; 1014 C; 958 G; 748 T; 0 other;

Alignment Scores:
 Pred. No.: 1,46e-294 Length: 3564
 Score: 3815.00 Matches: 727
 Percent Similarity: 99.86% Conservative: 0
 Best Local Similarity: 99.86% Mismatch: 1
 Query Match: 99.82% Indels: 0
 DB: 24 Gaps: 0

US-09-836-392-21 (1-728) x AAD33695 (1-3564)

QY 1 MetLeuArgHisLeuAlaThrAspAlaMetLysAsnPheSerGluPheArgGlnGlu 20
 DB 74 ATGCTGAGGACACCTCGGGCCACCGCATGCAAGAACTCTCCGAGTCCGCGAGAG 133
 QY 21 AlaSerMetLeuHisAlaLeuGlnHisProGlyIleValAlaLeuIleIleSerIle 40
 DB 134 GCCAGCATGCTCCAGCGCTGCGAGCACCCCTCATCTGTCGCTATCGCATGACGATC 193
 QY 41 HisProLeuCysPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
 DB 194 CACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253
 QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlyHisMetLeuThrGlnLysIleAla 80
 DB 254 GAGACGCGACGAGATCTCTTATACCCCTGAGGACACATGCTACCCAAAAATATAGCC 313
 QY 81 TyrGlnIleAlaSerGlyLeuAlaTrpLeuHisLysLysAsnIleIlePheCysAspLeu 100
 DB 314 TACCAAGATCGCTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 373
 QY 101 LysSerAspAsnIleLeuValIleTrpSerLeuAspValLysGluHisIleAsnIleLysLeu 120
 DB 374 AAGTCGACACAACTTGT 433

QY 121 SerAspTyrGlyIleSerArgGlnSerPheHisGluValAlaLeuGluValGluGlyThr 140
 |||||
 Db 434 TCGATGATACGGGATTTTCGAGGAGTCTATTCATGAGGCGCCCTTAGCGGTGGAGGACT 433
 QY 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluValAspMet 160
 |||||
 Db 494 CTGGCTACAGAGCCCGAGAGATCAGAGCCCTCGCATTTGATATGATGAGAGATGATG 553
 QY 161 PheSerTyrGlyMetValLeuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlyHis 180
 |||||
 Db 554 TTTCTTATGAGAAATGGTGTCTACAGAGTGTCTCAGAGACGCGCCCTGCACGCGGCGAC 613
 QY 181 HisGlnLeuGlnIleAlaValLysLysLeuSerLysGlyIleArgProValLeuGluGlnPro 200
 |||||
 Db 614 CACGACCTCCAGATTTCCCAAGAAAGCTGTCCCAAGGCGATCCGCGGTTCTGGGGCAGCG 673
 QY 201 GluGluValGlnPheArgArgLeuGlnAlaLeuMetGluGlySerTyrAspThrLysPro 220
 |||||
 Db 674 GAGGAGTGCAGATTCGGGGAGTGCAGGCGCTCATGTGAGTGGAGTGGGACACTAAGCCA 733
 QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
 |||||
 Db 734 GAGAAACGAGACGCTGGCCCTGTCTGGTGGAGCCAGATGAGAGACCGACTTTTGGCCACC 793
 QY 241 PheMetTyrGluLeuLeuLysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGln 260
 |||||
 Db 794 TTTATTTATGAACTGTGTCTGGAGAGACAGACGCTTTCTTCTCATCCAGGGCCAGAG 833
 QY 261 TyrThrValValPheThrAspGlyLysGluGluSerArgAsnTyrThrValValAsnThr 280
 |||||
 Db 854 TACACCGTGTGTGGTGGATGAAAGAGAGATCCAGCAATACAGCGGTGGAGACACA 913
 QY 281 GluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
 |||||
 Db 914 GAGAAAGGCTCTATGAGAGTGCAGAGATGTGTGCTGGAGTGAAGGAGACTCCAG 973
 QY 301 LeuGlnValGlnArgSerLeuThrAlaThrGluAspGlnLysIleTyrIleTyrThr 320
 |||||
 Db 974 CTCACAGTCCAGAGATCCCTGTGACAGACCGAGAGACCAAGAAATCTACATCTACAC 1033
 QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaVal 340
 |||||
 Db 1034 CTCAGAGGATGTGCTTAAACACACCCCAAGCGCTGTGATCTCCAGCTGTCTGTC 1093
 QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
 |||||
 Db 1094 ACTGCTTTCTGGCCCTGCTGCTTATTAAGAAATTCCTACTGCTTAAAGGCGCTC 1153
 QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
 |||||
 Db 1154 GCCGATGGCTGTGGCTGTGTTCCGTTGGTGGGCGACCCCAAGAGCACCTGCTCC 1213
 QY 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
 |||||
 Db 1214 TACCTGTGTCTACACACACACCAAGCTCCAAAGTTCAGATCCGGATGAGAGACGACG 1273
 QY 401 GlnAspProTyrProValLysAlaMetGluValValAsnSerGlySerGluValIleTyrPyr 420
 |||||
 Db 1274 CAGAACCCCTACCACTGAGAGCCATGAGAGTGTGACAGCGGCTGTGGGTGTATC 1333
 QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluLysArgArgLeu 440
 |||||
 Db 1334 AGCAATAGGCGCGGCTCTGTTCATCGACTGTGCTCCCTGAGATTCGACAGCGGCTG 1393
 QY 441 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGluGlyArgGly 460
 |||||
 Db 1394 GAGCCCTACATGCGCCCTCCCAAGTGTAGTCACTGTGTGAGGCGGCGGCGGAGG 1453
 QY 461 GluGluValValTyrPysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
 |||||
 Db 1454 GAGGAGGTGTCTGTGCTGTGATGACAAAGGCCCAACTCTTGTGTATGACCACTCCACC 1513
 QY 481 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 500

Db 1514 ACTTACACACTGTGTGCTCCCGGATCTTCTCGGGGTCCCAAGCCCTCAGGAGCATGTT 1573
 QY 501 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysVal 520
 |||||
 Db 1574 CCGGTGCGGCGCTTGGACAGCGAAACCCCGGCGACAGCCAGCCCAACCAAGAGTGG 1633
 QY 521 ProGluGlyAspSerIleAlaAspValSerIleMetTyrSerGluGluLeuGlyThrGln 540
 |||||
 Db 1634 CCGAGAGGGAGACTCCATCCCGGACGTGACATCATATACAGAGAGAGTGGGCGACG 1693
 QY 541 IleLeuIleHisGlnGlnLeuSerLeuThrAspTyrCysSerMetSerSerTyrSerSer 560
 |||||
 Db 1694 ATCTATATCCACAGGAATCAGTCACTGACTACTGCTCATATGCTCTTCTTCTCATCC 1753
 QY 561 ProProArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 580
 |||||
 Db 1754 CCACCCCGGCGAGCTGCGAGTCCCTCAAGCTCCCAAGCTCCCAAGCTCCCAAGCTTTC 1813
 QY 581 SerValProPheSerThrAspCysGluAspSerAspMetLeuHisThrProGlyAlaAla 600
 |||||
 Db 1814 AGTGTGCTTTCTCCACCACTGCGAGACTAGACATGCTACTATACGCCGTGTGTGCC 1873
 QY 601 SerAspArgSerGluHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 620
 |||||
 Db 1874 TCCGACAGTCTGAGCATACCTGACCCCAATGAGACGGGAGACCTTCAGCCAGCCTG 1933
 QY 621 GlnAlaValLysIleLeuAlaValArgAspLeuIleTyrPvalProArgArgGlyAsp 640
 |||||
 Db 1934 CAGGCGGTGAAGATCTCCCGGCGAGACATCATTTGGTGGTCCCGGCGGCGGTGAGAT 1993
 QY 641 ValIleValIleGlyLeuGluLysAspSerGluAlaGlnArgGlyArgValIleAlaVal 660
 |||||
 Db 1994 GTTATGCTATGTGCTGAGAGAGATTTGTGGCGCCACGGGCGGCGGATGTGCGCTG 2053
 QY 661 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 680
 |||||
 Db 2054 TTAAGACCCGAGAGCTGACTCCGCTAGGGGTGTGTGTGATGCTGCTGCGGCGCAAG 2113
 QY 681 AspThrValValCysThrPheGluAsnGluAsnThrGluTyrPysLeuAlaValTyrArg 700
 |||||
 Db 2114 GACACTGTGTGTGACCTTTGAAATGAAACACAGACTGTGTGCTGCGCTGTGGAG 2173
 QY 701 GlyTyrGlyAlaArgGluPheAspIlePheTyrGlnSerTyrGluGluLeuGlyArgLeu 720
 |||||
 Db 2174 GGTGTGGGCGCCAGGAGTTCGACATTTCTTACACATCTTACGAGAGCTGGCGGCTG 2233
 QY 721 GluAlaCysThrArgLysArgArg 728
 |||||
 Db 2234 GAGGCTTGACATCGCAAGAGAGAG 2257
 RESULT 4
 AAS72238
 ID AAS72238 standard; cDNA; 2652 BP.
 XX AAS72238;
 AC AAS72238;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #8042.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 OS
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.
XX (HSE-) HYSD INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG08051.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 1; SEQ ID No 8042; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridization probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2652 BP; 565 A; 824 C; 766 G; 497 T; 0 other;
XX
Alignment Scores:
Pred. No.: 8,64e-288 Length: 2652
Score: 3728.00 Matches: 719
Percent Similarity: 98.49% Conservative: 0
Best Local Similarity: 98.49% Mismatches: 8
Query Match: 97.54% Indels: 3
DB: 23 Gaps: 0
US-09-836-392-21 (1-728) x AAS72238 (1-2652)
QY 1 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnIlePheSerGluPheArgGlnGlu 20
DB 445 ATGCTGAGGACCTGGGGGACCGATGCCATGAAGAACTTCCGAGTCCGGCAGAG 504
QY 21 AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 40
DB 505 GCCAGCGTGTGACGCGCGTGCAGCACCCCTGCATCGTGGCGGCATCGCATGATC 564
QY 41 HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 60
DB 565 CACCGCTGTGCTTGGCCCTGGAGCTGGCGGCTGCAGCGCCCAACCGCGTCTGCC 624
QY 61 GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnIleValAla 80
DB 625 GAGAAAGCCAGATCTCTCTTATACCCCTGGGACATGCTCACCAAAAATAGCC 684
QY 81 TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 100
DB 685 TACCAAGATCGCTTGGCGCTGGCTTACCTGCACAAAGAAACATCACTCTGTGACCTG 744
QY 101 LysSerAspAsnIleLeuValAlaTyrSerLeuAspValLysGlnHisIleAsnIleLysLeu 120
DB 745 AAGTCGACAAACATTCCTGGTGTGCTCCCTTGACGTCAAGAGCACATCAATCAAGCTA 804

QY 121 SerAspTyrGlyIleSerArgGlnSerPheIleGluValAlaLeuGlnGlyThr 140
DB 805 TCTGACTACGGATTTCCAGAGCATCTTCATGAGAGGCGCCCTAGCGGGAGGCACT 864
QY 141 ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet 160
DB 865 CCGGCTACAGAGCCCGCAGAGATCAGAGCCCTCGCATTTGTATATGATGAAGATAGATG 924
QY 161 PheSerTyrGlyMetValLeuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlnHis 180
DB 925 TTCCTCATGATGATGTGCTGCTACGAGTTGCTCTCAGAGAGCGCCCTGCACGGCGCAC 984
QY 181 HisGlnLeuGlnIleAlaLysLysLeuSerGlyIleLeuArgProValLeuGlnGlnPro 200
DB 985 CACCACTCCAGATTCGCAAGAGGCTGTCCAAAGGCTCCGCGGTTCTGGGGGACGCC 1044
QY 201 GluGluValGlnPheArgArgLeuGlnAlaLeuMetMetGluCysTyrAspThrLysPro 220
DB 1045 GAGGAAGTGCAGTTCGGGCGACTGCAGCGCTCATGATGAGTGTGCGACATAAGCCA 1104
QY 221 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 240
DB 1105 GAGAGAGGACCGCTGGCGCTGCGGTGTGAGCCAGATGAAGGACCGACTTTGCCAC 1164
QY 241 PheMetTyrGluLeuCysGlyLysGlnThrAlaPheSerSerGlnGlnGln 260
DB 1165 TTCATGATGATACCTGCTGCTGGGAGCAGACAGCCCTTCTCATCCCGGCGCCAGAG 1224
QY 261 TyrThrValAlaPheThrAspGlyLysGlnLysArgAsnTyrThrValAlaAsnThr 280
DB 1225 TACACCGTGTGTTGGGAGTGAAGAGAGTCCGAGACTACACGGTGGTGAACACA 1284
QY 281 GluLysGlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGln 300
DB 1285 GAGAAAGCCCTCATGAGAGTGCAGAGATGTGCTGCCCTGGATGTAAGTACGTGCAG 1344
QY 301 LeuGlnValGlnArgSerLeuThrPheAlaThrGluAspGlnLysIleTyrThr 320
DB 1345 CTCAGAGTCCAGAGATCCCTGTGTGACAGCCAGCCAGACCAAGAAATCTACATCCACC 1404
QY 321 LeuLysGlyMetCysProLeuAsnThrProGlnAlaLeuAspThrProAlaValVal 340
DB 1405 CTCAGAGGCTGTGCGCCCTTAAACACACCCCAACAGGCTGTGGATCTCCAGCTGTGTC 1464
QY 341 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 360
DB 1465 ACCTGCTTCTGGCCGCGCTGTTATTAAGAAATCTCACTGCTTACGGCGGCTC 1524
QY 361 AlaAspGlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSer 380
DB 1525 GCCGATGGCTGTGTGCTGTGTTCCCGTGGTGGCGGACCCCAAGAGACAGCTGCTC 1584
QY 381 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 400
DB 1585 TACCTGTGCTCACACAGAGCAAGGTCCAAAGTTCAGCATCGGGTGAAGACGCGG 1644
QY 401 GluAsnProTyrProValLysAlaMetGluValValAsnSerGlySerGluValAlaTyr 420
DB 1645 CAGAACCCCTACCCAGAGGAAGCCATGAGAGTGTGTACAGCGGCTGTGAGGTCTGTAC 1704
QY 421 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgArgLeu 440
DB 1705 AGCAATGGCGCGGCGCTCTGTGATGAGTGTGCTCCCTGAGAGATCTCAGCGGCTG 1764
QY 441 GluProTyrMetAlaProSerMetValThrSerValValLysSerSerGluIleArgGly 460
DB 1765 GAGCCCTACATGGCCCTCTCATGTTACGTCAAGTGTGAGCTGTGAGGAGAGGG 1824
QY 461 GluGluValValTyrCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerThr 480
DB 1825 GAGAGAGTGTCTGTGCTGCTGATGACAAAGCCAACTCCTGGTGTGATGACTCCAC 1884

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:21:44 ; Search time 29 Seconds

(without alignments)
738.617 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822

Sequence: 1 MLRLRATDAKNSEPRFE.....IFYSYELGRLACTRRKR 728

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Issued Patents: AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	243.5	6.4	505 1	US-08-222-616-20 Sequence 20, Appl
2	243.5	6.4	505 2	US-08-446-648-20 Sequence 20, Appl
3	243.5	6.4	505 3	PCT-US95-04238-20 Sequence 20, Appl
4	243	6.4	455 3	US-09-221-235-5 Sequence 5, Appl
5	243	6.4	455 3	US-09-221-928-5 Sequence 5, Appl
6	243	6.4	455 3	US-09-221-527-5 Sequence 5, Appl
7	243	6.4	455 4	US-09-221-236-5 Sequence 5, Appl
8	243	6.4	455 4	US-09-221-416-5 Sequence 5, Appl
9	243	6.4	455 4	US-09-221-245-5 Sequence 5, Appl
10	243	6.4	455 4	US-09-163-115-5 Sequence 5, Appl
11	243	6.4	455 4	US-09-221-528-5 Sequence 5, Appl
12	243	6.4	455 4	US-09-593-553-5 Sequence 5, Appl
13	243	6.4	455 4	US-09-221-237-5 Sequence 5, Appl
14	241.5	6.3	506 4	US-08-426-509A-6 Sequence 6, Appl
15	241.5	6.3	511 5	PCT-US95-05008-6 Sequence 6, Appl
16	238	6.2	800 2	US-08-469-537A-72 Sequence 72, Appl
17	238	6.2	800 2	US-08-469-537A-78 Sequence 78, Appl
18	238	6.2	937 2	US-08-469-537A-105 Sequence 105, Appl
19	236	6.2	263 3	US-09-035-706-5 Sequence 5, Appl
20	236	6.2	263 3	US-08-955-841-5 Sequence 5, Appl
21	236	6.2	263 4	US-09-390-425-5 Sequence 5, Appl
22	236	6.2	263 4	US-09-566-906-5 Sequence 5, Appl
23	236	6.2	821 1	US-07-928-464-2 Sequence 2, Appl
24	236	6.2	821 1	US-08-003-311B-2 Sequence 2, Appl
25	236	6.2	821 1	US-08-261-432-2 Sequence 2, Appl
26	236	6.2	821 5	PCT-US93-07347-2 Sequence 2, Appl
27	233	6.1	793 3	US-09-031-563-23 Sequence 23, Appl

28	233	6.1	795 4	US-09-392-277-23 Sequence 23, Appl
29	228.5	6.0	579 4	US-09-529-279-4 Sequence 4, Appl
30	228.5	6.0	590 4	US-09-529-279-15 Sequence 15, Appl
31	228	6.0	296 2	US-08-701-191A-29 Sequence 29, Appl
32	228	6.0	668 1	US-08-205-018-2 Sequence 2, Appl
33	228	6.0	859 1	US-08-395-580-2 Sequence 2, Appl
34	228	6.0	859 5	PCT-US95-02192-2 Sequence 2, Appl
35	224.5	5.9	943 2	US-08-469-537A-107 Sequence 107, Appl
36	223.5	5.8	1584 4	US-08-457-040B-27 Sequence 27, Appl
37	219.5	5.7	630 2	US-08-596-319-2 Sequence 2, Appl
38	218	5.7	283 2	US-08-701-191A-41 Sequence 41, Appl
39	217.5	5.7	262 2	US-07-857-224B-57 Sequence 57, Appl
40	216.5	5.7	968 4	US-09-228-966-76 Sequence 76, Appl
41	215	5.6	306 2	US-08-701-191A-12 Sequence 12, Appl
42	215	5.6	306 4	US-09-457-040B-40 Sequence 40, Appl
43	215	5.6	374 4	US-08-857-076-108 Sequence 108, Appl
44	215	5.6	1382 2	US-08-737-715-2 Sequence 2, Appl
45	215	5.6	1382 4	US-09-457-040B-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-222-616-20
Sequence 20, Application US/08222616
Patent No. 5635177

GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESS: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,616
FILING DATE: 4-APR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00586
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/826935
FILING DATE: 22-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 821P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: 1linear
US-08-222-616-20

Query Match 6.48; Score 243.5; DB 1; Length 505;
Best Local Similarity 30.58; Pred. No. 1.3e-14;
Matches 75; Conservative 49; Mismatches 77; Indels 45; Gaps 11;

QY 15 SEFROASMLHALOHPCTIVALIGISI--HPLCFALFLAPLSSLNTVSENARDSSFFPLG 72
DB 272 NDFLEAQMKNLHNPILQIYAVCTLEDPIYITELMRHSLQEYL-QNDTGSKI----- 326
QY 73 HMLTOKI--AYOIASGLAYLHKNNITFDLKSNDILWMSLDVKEHINKISDYGISR--- 127
DB 327 -HLTQOYDMAQVASGMYLESNNYIHRDLAARNVL-----VGEHNITYKVADEFLARVFK 380
QY 128 -----OSFEHGAAGVEGTPGYOAPRIPIYDEKVDMSYGMVLEYELLS-GORPALG 179
DB 381 VDNEIDYESRHEIKLPYKWT-----APEAIRSNKFSIKSDVWSFGILLYEITTYGKMPYSG 436
QY 180 HHQLQIAKLSKSGIRPYLGOPEEYQFRRLQALMECDTKREKRLALSVSQMKDPTFA 239
DB 437 MTGAQVLOMLAQNRY--LPQPSNCP-QQFYNIMLECNNAEPRK-----PTFE 481
QY 240 TFMTEL 245
DB 482 TLRWKL 487

RESULT 2

US-08-446-648-20
Sequence 20, Application US/08446648
Patent No. 6331302

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Mapatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,648
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.

REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear
US-08-446-648-20

Query Match 6.48; Score 243.5; DB 4; Length 505;
Best Local Similarity 30.58; Pred. No. 1.3e-14;
Matches 75; Conservative 49; Mismatches 77; Indels 45; Gaps 11;

QY 15 SEFROASMLHALOHPCTIVALIGISI--HPLCFALFLAPLSSLNTVSENARDSSFFPLG 72
DB 272 NDFLEAQMKNLHNPILQIYAVCTLEDPIYITELMRHSLQEYL-QNDTGSKI----- 326
QY 73 HMLTOKI--AYOIASGLAYLHKNNITFDLKSNDILWMSLDVKEHINKISDYGISR--- 127
DB 327 -HLTQOYDMAQVASGMYLESNNYIHRDLAARNVL-----VGEHNITYKVADEFLARVFK 380
QY 128 -----OSFEHGAAGVEGTPGYOAPRIPIYDEKVDMSYGMVLEYELLS-GORPALG 179
DB 381 VDNEIDYESRHEIKLPYKWT-----APEAIRSNKFSIKSDVWSFGILLYEITTYGKMPYSG 436
QY 180 HHQLQIAKLSKSGIRPYLGOPEEYQFRRLQALMECDTKREKRLALSVSQMKDPTFA 239
DB 437 MTGAQVLOMLAQNRY--LPQPSNCP-QQFYNIMLECNNAEPRK-----PTFE 481
QY 240 TFMTEL 245
DB 482 TLRWKL 487

RESULT 3

PCT-US95-04228-20
Sequence 20, Application PC/TUS9504228

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee

REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids

TYPE: amino acid

TOPOLOGY: linear
PCT-US95-04228-20

Query Match 6.4%; Score 243.5; DB 5; Length 505;
 Best Local Similarity 30.5%; Pred. No. 1.3e-14;
 Matches 75; Conservative 49; Mismatches 77; Indels 45; Gaps 11;

15 SEFQASMLHALQPCIVALLIGISI--HPLCFALPLASSLNTVSENRDSSFIPLG 72
 272 NDFLEADQIMKRLRPRKILQIAVACTLEDPYIITELMRHGSLOEYL-QNDGSKRI----- 326
 73 HMLTQRT--AVQASGLAYLHKNNIIFCDKSDNLLWSLDVKEHINIKLSYGSIR--- 127
 327 -HLTQVMAAQAQVAGSMAYLESRNTHRDLAARNVL---VGEHNITKVAQDFGLARVFK 380
 128 -----OSFHEGALGVEGTPGYOAPETIRPIYDEKVMESYGMVLYELLS-GORPALG 179
 381 VDNEDVIESRHRKIKLPVMT-----APEARSAKFSIKSDVWSFGILITITIGKMPISG 436
 180 HHQLOIARKLSKIRPVLAGQPEEVOFRRLQALMECWDTPKPRPLASVSQMKDPFA 239
 437 MTGAQVIOQLAQNYR--LPQPSNCP-QQPYNIMLEQNNAPKPER-----PPFE 481
 240 TFMTELC 245
 482 TLRMTL 487

RESULT 4

US-09-221-235-5
 ; Sequence 5, Application US/09221235
 ; Patent No. 6043040
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNT-050
 ; CURRENT APPLICATION NUMBER: US/09/221,235
 ; CURRENT FILING DATE: 1998-12-28
 ; EARLIER APPLICATION NUMBER: 09/163,115
 ; EARLIER FILING DATE:
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 455
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-221-235-5

Query Match 6.4%; Score 243; DB 3; Length 455;
 Best Local Similarity 27.9%; Pred. No. 1.2e-14;
 Matches 69; Conservative 52; Mismatches 100; Indels 26; Gaps 7;

10 AKNFESEFROASMLHALQPCIVALLIGISIHPLCFAL--ELAPLSLNTVSENRDSS 67
 43 AVKLIKIEKEAELISVLSHRNITQFYGVLEPPNYGIVTEYASLSGLDYINSNR--SE 100
 68 FIDPLGHMLTOKIAYQIASGLAYLHKR---NIIFCDKSDNLLWSLDVKEHINIKLSYDG 124
 101 EMDMDHIMTW--ATDVAKGMHTLHMEAPYKVIHRDLKSRNVYIADGV-----LKICDFG 153
 125 ISROSEFHEGALGVEGTPGYOAPETIRPIYDEKVMESYGMVLYELLSGORPALGHQLO 184
 154 ASRFHNHTHMSLVGTFFPMARPEVIOSLPVSCTCTYSGVLYMELTREFVFKGLEGLQ 213
 185 IAKKLSKIRPVLAGQPEEVOF-----RRLQALMECWDTPKPRPLASVSQMKDPFA 239
 214 VAWL-----VYKNERLITPSSCPSPFALLHOCWEADAKRRSPFKQIISILSMSND 266
 240 TFMTELC 246
 267 TSLPDKC 273

RESULT 5
 US-09-221-928-5
 ; Sequence 5, Application US/09221928

Patent No. 6121030
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNT-050
 ; CURRENT APPLICATION NUMBER: US/09/221,928
 ; CURRENT FILING DATE: 1998-12-28
 ; EARLIER APPLICATION NUMBER: 09/163,115
 ; EARLIER FILING DATE:
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 455
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-221-928-5

Query Match 6.4%; Score 243; DB 3; Length 455;
 Best Local Similarity 27.9%; Pred. No. 1.2e-14;
 Matches 69; Conservative 52; Mismatches 100; Indels 26; Gaps 7;

10 AKNFESEFROASMLHALQPCIVALLIGISIHPLCFAL--ELAPLSLNTVSENRDSS 67
 43 AVKLIKIEKEAELISVLSHRNITQFYGVLEPPNYGIVTEYASLSGLDYINSNR--SE 100
 68 FIDPLGHMLTOKIAYQIASGLAYLHKR---NIIFCDKSDNLLWSLDVKEHINIKLSYDG 124
 101 EMDMDHIMTW--ATDVAKGMHTLHMEAPYKVIHRDLKSRNVYIADGV-----LKICDFG 153
 125 ISROSEFHEGALGVEGTPGYOAPETIRPIYDEKVMESYGMVLYELLSGORPALGHQLO 184
 154 ASRFHNHTHMSLVGTFFPMARPEVIOSLPVSCTCTYSGVLYMELTREFVFKGLEGLQ 213
 185 IAKKLSKIRPVLAGQPEEVOF-----RRLQALMECWDTPKPRPLASVSQMKDPFA 239
 214 VAWL-----VYKNERLITPSSCPSPFALLHOCWEADAKRRSPFKQIISILSMSND 266
 240 TFMTELC 246
 267 TSLPDKC 273

RESULT 6

US-09-221-527-5
 ; Sequence 5, Application US/09221527
 ; Patent No. 6146832
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNT-050
 ; CURRENT APPLICATION NUMBER: US/09/221,527
 ; CURRENT FILING DATE: 1998-12-28
 ; EARLIER APPLICATION NUMBER: 09/163,145
 ; EARLIER FILING DATE:
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 455
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-221-527-5

Query Match 6.4%; Score 243; DB 4; Length 455;
 Best Local Similarity 27.9%; Pred. No. 1.2e-14;
 Matches 69; Conservative 52; Mismatches 100; Indels 26; Gaps 7;

10 AKNFESEFROASMLHALQPCIVALLIGISIHPLCFAL--ELAPLSLNTVSENRDSS 67
 43 AVKLIKIEKEAELISVLSHRNITQFYGVLEPPNYGIVTEYASLSGLDYINSNR--SE 100
 68 FIDPLGHMLTOKIAYQIASGLAYLHKR---NIIFCDKSDNLLWSLDVKEHINIKLSYDG 124
 101 EMDMDHIMTW--ATDVAKGMHTLHMEAPYKVIHRDLKSRNVYIADGV-----LKICDFG 153

QY 125 ISROSFHEGALGVEGTGPGVQAEIRPRIYDEKVMFSYGVLYELLSGORPALGHQ 184
DB 154 ASRFNHTHTMSLVGTFFPMMAPEVIOSLPVSECTDYSGYVLMELTREVPEFKGLEQ 213
QY 185 IAKKLSKGRPVLAGPBEVQF-----RRQALMECWDTPKPKRPLALSVSOMKDPFA 239
DB 214 VAML-----VVEKNERLTIPSSCPRSFAELHOCWEADAKKRPSFOIISLESMSND 266
QY 240 TFWYELC 246
DB 267 TSLPDKC 273

RESULT 7
US-09-221-236-5
Sequence 5, Application US/09221236-
Patent No. 6146841
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,236
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-236-5

Query Match 6.4%; Score 243; DB 4; Length 455;
Best Local Similarity 27.9%; Pred. No. 1,2e-14;
Matches 69; Conservative 52; Mismatches 100; Indels 26; Gaps 7;

QY 10 AMKNSEFEQASMLHALQHPICVIALIGISHPICFAL--ELAPLSLNTVISENARDSS 67
DB 43 AVKLLKTEKEAELISVSHRNIIQFYGVLEPPYGIETEVASIGSLDYINSNR--SE 100
QY 68 FIPGLHMLTOKIAYOIASGLAYLHKK---NIIFCDLKSNDILWLSLDVKEHINIKLSDYG 124
DB 101 EMDMDHIMTW--ATDVAGMHTLHNEAPVKYIHRDLKSRNVIAADGV-----LKICDFG 153
QY 125 ISROSFHEGALGVEGTGPGVQAEIRPRIYDEKVMFSYGVLYELLSGORPALGHQ 184
DB 154 ASRFNHTHTMSLVGTFFPMMAPEVIOSLPVSECTDYSGYVLMELTREVPEFKGLEQ 213
QY 185 IAKKLSKGRPVLAGPBEVQF-----RRQALMECWDTPKPKRPLALSVSOMKDPFA 239
DB 214 VAML-----VVEKNERLTIPSSCPRSFAELHOCWEADAKKRPSFOIISLESMSND 266
QY 240 TFWYELC 246
DB 267 TSLPDKC 273

RESULT 8
US-09-221-416-5
Sequence 5, Application US/09221416
Patent No. 6153417
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,416
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-416-5

Query Match 6.4%; Score 243; DB 4; Length 455;
Best Local Similarity 27.9%; Pred. No. 1,2e-14;
Matches 69; Conservative 52; Mismatches 100; Indels 26; Gaps 7;

QY 10 AMKNSEFEQASMLHALQHPICVIALIGISHPICFAL--ELAPLSLNTVISENARDSS 67
DB 43 AVKLLKTEKEAELISVSHRNIIQFYGVLEPPYGIETEVASIGSLDYINSNR--SE 100
QY 68 FIPGLHMLTOKIAYOIASGLAYLHKK---NIIFCDLKSNDILWLSLDVKEHINIKLSDYG 124
DB 101 EMDMDHIMTW--ATDVAGMHTLHNEAPVKYIHRDLKSRNVIAADGV-----LKICDFG 153
QY 125 ISROSFHEGALGVEGTGPGVQAEIRPRIYDEKVMFSYGVLYELLSGORPALGHQ 184
DB 154 ASRFNHTHTMSLVGTFFPMMAPEVIOSLPVSECTDYSGYVLMELTREVPEFKGLEQ 213
QY 185 IAKKLSKGRPVLAGPBEVQF-----RRQALMECWDTPKPKRPLALSVSOMKDPFA 239
DB 214 VAML-----VVEKNERLTIPSSCPRSFAELHOCWEADAKKRPSFOIISLESMSND 266
QY 240 TFWYELC 246
DB 267 TSLPDKC 273

RESULT 9
US-09-221-245-5
Sequence 5, Application US/09221245
Patent No. 6180358
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,245
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: US 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-245-5

Query Match 6.4%; Score 243; DB 4; Length 455;
Best Local Similarity 27.9%; Pred. No. 1,2e-14;
Matches 69; Conservative 52; Mismatches 100; Indels 26; Gaps 7;

QY 10 AMKNSEFEQASMLHALQHPICVIALIGISHPICFAL--ELAPLSLNTVISENARDSS 67
DB 43 AVKLLKTEKEAELISVSHRNIIQFYGVLEPPYGIETEVASIGSLDYINSNR--SE 100
QY 68 FIPGLHMLTOKIAYOIASGLAYLHKK---NIIFCDLKSNDILWLSLDVKEHINIKLSDYG 124
DB 101 EMDMDHIMTW--ATDVAGMHTLHNEAPVKYIHRDLKSRNVIAADGV-----LKICDFG 153
QY 125 ISROSFHEGALGVEGTGPGVQAEIRPRIYDEKVMFSYGVLYELLSGORPALGHQ 184
DB 154 ASRFNHTHTMSLVGTFFPMMAPEVIOSLPVSECTDYSGYVLMELTREVPEFKGLEQ 213
QY 185 IAKKLSKGRPVLAGPBEVQF-----RRQALMECWDTPKPKRPLALSVSOMKDPFA 239
DB 214 VAML-----VVEKNERLTIPSSCPRSFAELHOCWEADAKKRPSFOIISLESMSND 266
QY 240 TFWYELC 246

RESULT 15
PCT-US95-05008-6
Sequence 6, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
KINASES
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 511 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-05008-6

Query Match 6.3%; Score 241.5; DB 5; Length 511;
Best Local Similarity 30.2%; Pred. No. 2,1e-14;
Matches 74; Conservative 49; Mismatches 79; Indels 43; Gaps 11;

QY 15 SEFROEASMLHALQHPCTIVALIGISI--HPLCFALIELAPLSSLNTVLSENARDSSFIPLG 72
DB 278 NDFLEAQIMENLRHPKLIQLYAVCTLEDPIYITIELMRHGSLOEYL-QNDTGSRI----- 332
QY 73 HMLFO-KIAYQIASGLAVLHKNNIIFCDLKSNDNIIVMSLDVKEHINIKLSDYGISR---- 127
DB 333 HLTQOYDMAAQVAGSMAYLESRNTHRDLAARNVL-----VGEHNTRYKVADEGLARVERV 387
QY 128 -----OSFHEGALVESTPGYQAEIRPRIYDEKVMFSYGMVLYELLG-GORPALGH 180
DB 388 DNEDIYESRHRHKLPLKWT-----APEAIRSNKFSIKSDVMSFGIILYIITIGKMPYSGM 443
QY 181 HOLQIAKRLSKIRPVLAQPEEVQFRRLQALMECWDTRPEKRPALSVSOMKDPTEFAT 240
DB 444 TGAQVITQMLAQNYR--LPQPSNCP-QQFYNIMTECMNAEPKER-----PTFET 488
QY 241 FMYEL 245
DB 489 LRWKL 493

Search completed: April 8, 2003, 10:26:57
Job time : 37 secs

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:21:03 ; Search time 49 Seconds
(without alignments)
1428.283 Million cell updates/sec

Title: us-09-836-392-21

Perfect score: 3822
Sequence: 1 MLRLRLATDANKNFSEKRFQ.....IFQSYELGRLEACTRRRR 728

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	402	10.5	1090	2 T33475	hypothetical prote
2	315	8.2	412	2 T10671	protein kinase hom
3	314	8.2	407	2 G84635	probable protein k
4	285	7.5	954	1 S68178	mixed-lineage prot
5	278.5	7.3	546	2 D84555	probable protein k
6	273.5	7.2	394	2 U00229	mixed-lineage prot
7	270.5	7.1	553	2 T04683	hypothetical prote
8	270.5	7.1	1171	2 T12956	hypothetical prote
9	268.5	7.0	438	2 C86273	protein kinases ho
10	268	7.0	475	2 T12955	probable protein k
11	267.5	7.0	1147	2 F66297	hypothetical prote
12	263	6.9	1029	2 H86179	hypothetical prote
13	263	6.9	1338	2 T18287	protein-tyrosine k
14	260.5	6.8	847	1 A53800	mixed-lineage prot
15	260.5	6.8	1248	1 B96827	hypothetical prote
16	259.5	6.8	545	2 T05675	hypothetical prote
17	257	6.7	410	2 B35670	protein-tyrosine k
18	254.5	6.7	506	1 S24553	protein-tyrosine k
19	251	6.6	736	2 T05137	protein kinase hom
20	249	6.5	406	2 T52626	probable mitogen-a
21	249	6.5	1015	2 T00726	probable serine/th
22	248.5	6.5	462	2 T02726	probable protein k
23	248.5	6.5	512	2 T48552	protein-tyrosine k
24	247	6.5	1257	2 T00486	serine/threonine-s
25	245.5	6.4	392	2 T47684	serine/threonine-s
26	243.5	6.4	505	2 T138396	protein-tyrosine k
27	242.5	6.3	314	2 T16052	hypothetical prote
28	242.5	6.3	879	2 T02728	serine/threonine-s
29	241	6.3	982	2 T06576	probable protein k

30	240.5	6.3	364	2 G71410	probable protein k
31	239	6.3	390	2 T01451	protein kinase hom
32	238.5	6.2	356	2 T50811	ser/thr specific p
33	238.5	6.2	402	2 T51791	ser/thr specific p
34	238.5	6.2	412	2 T02725	probable serine/th
35	238.5	6.2	509	2 T04688	hypothetical prote
36	238	6.2	937	2 A45082	neurotrophic recep
37	238	6.2	1303	1 S27396	phytochrome / prot
38	236.5	6.2	453	2 T19663	tyrosine kinase (f
39	236.5	6.2	497	1 S43532	protein-tyrosine k
40	236	6.2	395	2 T00574	probable protein k
41	236	6.2	821	2 T48400	serine/threonine-p
42	235.5	6.2	323	2 S04328	protein-tyrosine k
43	235.5	6.2	738	2 F96701	hypothetical prote
44	235	6.1	391	2 T48115	protein kinase ATM
45	235	6.1	505	1 S24550	protein-tyrosine k

ALIGNMENTS

RESULT 1					
T33475					
hypothetical protein T27C10.5 - Caenorhabditis elegans					
C:Species: Caenorhabditis elegans					
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999					
C:Accession: T33475					
R:Zhu, H.J.; Graves, T.; Hawkins, M.					
Submitted to: The EMBL Data Library, October 1998					
A:Description: The sequence of C. elegans cosmid T27C10.					
A:Reference number: Z21354					
A:Accession: T33475					
A:Status: Preliminary; translated from GB/EMBL/DBJ					
A:Molecule type: DNA					
A:Residues: 1-1090 <ZHU>					
A:Cross-references: EMBL:AF098504; PIDN:AAC67413.1; GSPDB:GN00019; CESP:T27C10.5					
A:Experimental source: strain Bristol N2; clone T27C10					
C:Genetics:					
A:Gene: CESP:T27C10.5					
A:Map position: 1					
A:Introns: 182/3; 260/2; 319/2; 796/3; 826/3; 979/3; 1038/2					
Query Match					
Best Local Similarity 10.5%; Score 402; DB 2; Length 1090;					
Matches 123; Conservative 86; Mismatches 175; Indels 60; Gaps 19;					
QY	10	AKKNFSEFEQASMLHALDHPICVALIGISHPCLAEPLASLNTVLSERNADSPRI	69		
DB	457	ACRAYCTSQEELSLSRKHPNIVIGVCTPPLSLVELAPLGAINGLGHRRAGTYL	516		
QY	70	PLGHMLTKRIAYQIASGLAYLHKRNITFCDLKSDNITVSLDV---KEHINIKLSYGI	125		
DB	517	SLG--VIRKSAVQVAVRALIYLSAHITLYRDKSENVLGWRFAPFSPQDVILKLDYGI	574		
QY	126	SQSFHE--GALVEGEPGQAPRI---RRIVYDEVVDMEFSGMVLXELLSQRPALGHN	181		
DB	575	SSSVLPSSGAKGFGGTGGWAPRIIVFNSEETQVDDCFSGMFLYELTLKFP--RS	632		
QY	182	QIQIAIKLSKGRIPVVGQEEVOFRRLQALMMECWPTKEKRPALSLVSOMKDPFFATF	241		
DB	633	EEIVKRMMDGARPYVLLPHELLPTPMDDLIVHCWNAHESRSSQLVGFCAPEFTH	691		
QY	242	MYELC-CGR---QTAFS-----SQGGRYVVFVDGKEESNRYVNV-TEKGLMEVOR	289		
DB	692	LIDVCEIAGEALPPTQLMAVGITDEIDDPDFEAQLV---LSGRENVMGCTGYFVDKRS	748		
QY	290	MCCP--GMVSOQLQVQBSLMTATEDQKIYITLKMCPLNTPQQLDPAVVTCTFLAVP	347		
DB	749	IELPHRGKYV---KYRDSVSCDECGQVTVY---GISLHETGH--LQPLSLNGTLICAP	800		
QY	348	VIKNSYVLIAGLADLVAVFPVVRGTPKDSCSYLCSHTANRKSFIADEDARQNPYPK	407		
DB	801	ELISNDVYLIT--ISDKQIVLKL---SESNSVSHL-----GITDSYEIR	840		

QY 408 AMEVNNGS-EWYNSNGPLVYI 429
 DB 841 TATFLGNGSTROIWAGHSEGRISI 864

RESULT 2

T10671
 protein kinase homolog F6E21.90 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000

C:Accession: T10671
 R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro

submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16533

A:Accession: T10671
 A:Molecule type: DNA

A:Residues: 1-412 <BEV>
 A:Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90

A:Experimental source: cultivar Columbia; BAC clone F6E21
 C:Genetics:

A:Gene: ATSP:F6E21.90
 A:Map position: 4

A:Introns: 300/2
 C:Superfamily: kinase-related transforming protein; protein kinase homology

F:135-392/Domain: protein kinase homology <KIN>

Query Match

Best Local Similarity 8.2%; Score 315; DB 2; Length 412;
 Matches 82; Conservative 44; Mismatches 73; Indels 32; Gaps 9;

QY 16 EFRQNSMLHALQPCIVALLIGISHP--CFALFLAPLSSLTNTVLSNARDSSFIPLGH 73

DB 176 OFQOEVSMLANKHPNIVFIRIGACIRPMWCITVEYAKGSVOFLTK--RONRAVPLK 233

QY 74 MLNOKIAYQASGLAYLHKNIIFCDLKSNDILWMSLDVKEHINIKSDIGISROSH-E 132

DB 234 AVMO--ALDVARGMATVHEHNFIRHDKSDNLI--SADK--RSIKIADGVAHIEVOTE 286

QY 133 GAGVGTEPGYQAPETIRPRIVDEKYMFSYGMVLYELLSGORPALGHQLOIA-KLISK 191

DB 287 GMPREGTGRMAPEMIGHNRYQKVDYVSGIVLWELITGLPFGMNTAVQAFAFVNR 346

QY 192 GIR-----PYLQPEYVQFRRLQALMECDTKPEKPLALSVSOKM 234

DB 347 GVRPTVPADCLPVLGE-----IMTRCWDADPEVAPCFAEIYNLLE 386

RESULT 3

G84635
 probable protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002

C:Accession: G84635
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Sha, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanNaken, S.E.; Umayam, L.; Tallon, L.

enus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 407, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MID:20083487; PMID:10617197
 A:Accession: G84635

A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-407 <STO>
 A:Cross-references: GB:AE002093; NID:g4337195; PIDN:AAD18109.1; GSPDB:GN00139

C:Genetics:
 A:Gene: AT2g24360
 A:Map position: 2

C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 8.2%; Score 314; DB 2; Length 407;
 Best Local Similarity 36.8%; Pred. No. 3e-11;
 Matches 82; Conservative 42; Mismatches 83; Indels 16; Gaps 9;

QY 16 EFRQNSMLHALQPCIVALLIGISHP--CFALFLAPLSSLTNTVLSNARDSSFIPLGH 73
 DB 171 OFQOEVSMLANKHPNIVFIRIGACIRPMWCITVEYAKGSVOFLTK--RONRAVPL-- 226

QY 74 MLNOKIAYQASGLAYLHKNIIFCDLKSNDILWMSLDVKEHINIKSDIGISROSH-E 132

DB 227 KLAVALDVARGMATVHEHNFIRHDKSDNLI--SADK--RSIKIADGVAHIEVOTE 281

QY 133 GAGVGTEPGYQAPETIRPRIVDEKYMFSYGMVLYELLSGORPALGHQLOIA-KLISK 191

DB 282 GMPREGTGRMAPEMIGHNRYQKVDYVSGIVLWELITGLPFGMNTAVQAFAFVNR 341

QY 192 GIRPVLQPEYVQFRRLQALMECDTKPEKPLALSVSOKM 234

DB 342 GVRPTV--PNDCLPVLSDIMTRCWDADPEVAPCFAEIYNLLE 381

RESULT 4

S68178
 mixed-lineage protein kinase 2 (EC 2.7.1.-) - human

C:Species: Homo sapiens (man)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 11-Jun-1999

C:Accession: S68178; I38044; S32468
 R:Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; S.

Eur. J. Biochem. 234, 492-500, 1995
 A:Title: Complete nucleotide sequence, expression, and chromosomal localisation of h

A:Reference number: S68178; MID:96128179; PMID:8536694
 A:Accession: S68178

A:Molecule type: mRNA
 A:Residues: 1-954 <DOR>

A:Cross-references: EMBL:X90846; NID:971419; PIDN:CAA63351.1; PID:9711420
 R:Kato, M.; Hirai, M.; Sugimura, T.; Terada, M.

Oncogene 10, 1447-1451, 1995
 A:Title: Cloning and characterization of MST, a novel (putative) serine/threonine ki

A:Reference number: I38044; MID:95249256; PMID:7731697
 A:Accession: I38044

A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA

A:Residues: 1-461; 'A', 'V', '465-470', 'S', '472-806', 'R', '808-817', 'A', '819-954 <RES>
 A:Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA8531.1; PID:g758593

R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kreuser, T.
 Eur. J. Biochem. 213, 701-710, 1993

A:Title: Identification of a new family of human epithelial protein kinases containi
 A:Reference number: S32467; MID:9338756; PMID:8477742

A:Accession: S32468
 A:Molecule type: mRNA

A:Residues: 244-464; 'AQAAGRORPAPML' <DO2>
 C:Genetics:

A:Gene: GDB:MK2; GDB:MST
 A:Cross-references: GDB:362654; GDB:624810; OMIM:600137

A:Map position: 19q13.1-19q13.2
 C:Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology

C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific prote
 F:23-76/Domain: SH3 homology <SH3>

F:76-364/Domain: protein kinase homology <KIN>
 F:104-112/Region: protein kinase ATP-binding motif

F:384-405/Region: leucine zipper motif
 F:419-440/Region: leucine zipper motif

F:449-463/Region: basic
 F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 7.5%; Score 285; DB 1; Length 954;
 Best Local Similarity 22.8%; Pred. No. 4.1e-09;

Matches 170; Conservative 93; Mismatches 275; Indels 208; Gaps 31;

QY 19 QEASMLHALQPCIVALLIGISHP--LCFALFLAPLSSLTNTVLSNARDSSFIPLGHMLT 76

DB 144 QEARLFGALQHPNIALRGACLNPPHICLVEYAKGALSRLVLAGR----VP-PHVLV 197

QY 77 OKIAYQASGLAYLHKNIIFCDLKSNDILWMSLDVKEHINIKSDIGISROSH-E 128

DB 198 -NNAVQVARGMNTLHNDAPVPIIHRDKSINILI--LEAIENHNLADTVLKITDPLGLARE 254

QY 129 SFHEGALGVESTPGYQAPETPRIVYDEKVDMESTGNTVLLSGORPALGHOLQIA-- 186
 DB 255 WKHTKSAAGTAYAMAPETPRIVYDEKVDMESTGNTVLLSGORPALGHOLQIA-- 314
 QY 187 ---KLLKSGIRPVLGQPEEVEFRRLQALMECMTKPEKRP-----LALSVSQ 232
 DB 315 VAMKRLPLPTSTPEP---FAR---LLECKMDPDHGRDFESSILKRELVIEOSALFQ 367
 QY 233 MKDPTF-----ATPMYELCCGKQTAFSSQGEYTVVMDKEE--- 271
 DB 368 MPLESPFSLQDMKLEIQHMFDDLTKEKELRSREBELLRQAQORQEDQLRRREDELA 427
 QY 272 SRATVYNTKGLMEVQRMCCPGKVCSCOLOVQSLTATDOKIYITLKG-----MC 325
 DB 428 EREMDIVERELHL-----MCOI-----SQEKPRVKKRKNFKRSRL 465
 QY 326 PLNTPQALDPAVYTCFLAY---PVI---KKNSTYVLGLADGLVAVPVRKTPKDCS 380
 DB 466 KLRGGSHISLPSPFKKITVQASPTLDKRGSDGASPPASPIIPLRAIRLIPVD--- 522
 QY 381 YLCSHTANRSKFSIADEDARQNPYPVK-----AMEV 412
 DB 523 ---CGSSSSSGSSGGSGTMSRGP--PKKELVGKKKGRMGPSSTLQKERVGGEERLKL 579
 QY 413 NSGEWYTSNPGI-----LVIDCASLEICRL-----EPTMASMTVS 452
 DB 580 GEGSKQWSSAPNLGKSPKHTPIAPGFASLMEFEAEDEGGSSVPSPYSTSYLSVP 639
 QY 453 VCSSEGRGEVEVWCLDKANSLVMYHSTYQL-----CA----- 486
 DB 640 LPAPSPCARAPWERTPAPRANGHARRCDLALLGALLGAVGLADYAEARAD 699
 QY 487 ---RYFCVPSPLDMFPVRPLDTEPPASHTANPKYEGSDIADVSIMYSEELQIL 542
 DB 700 EEOQRMLDGLFFPRAGRP--RGL--SPAPRHGREDEVPGLGLAP-----SATL 746
 QY 543 IHQSLNDYCSMSYSSSSPPROA--RSPSLPSSPASSSSVPFSP---DCEDSDMLTP 597
 DB 747 VLSLSSVSCNSTRLNSDSOEAPAPSPSPAPPTPSPSTNPLVDELESFKKP 806
 QY 598 GAADPRSE-----HDLTPMDG 613
 DB 807 GQSLPTHTVAVCAVSRGHRRTPSDG 832
 RESULT 5
 D84555
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: D84555
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; VanKlee, S.E.; Umayam, L.; Tallon, L.;
 Euss, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402:761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: D84555
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-546 <STO>
 A:Cross-references: GB:AE002093; NID:96598802; PIDN:AA80785.2; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g17700
 A:Map position: 2
 Query Match
 Best Local Similarity 7.38; Score 278.5; DB 2; Length 546;
 Matches 71; Conservative 45; Mismatches 90; Indels 17; Gaps 8;
 QY 16 EFROASMLHMQHCYIALGISIHP--PLCFALPLPLSSINTVLSNARDSFTPLGH 73

DB 327 EFSEGFIMKRVKKNVVOFLGACTRSPTLCITVEFARGSTIDFLH---KQCAFKLOT 383
 QY 74 MLTKIAYQIALSLAYLHKNIITCDLKSNDILWMSLDVKHINIKSDGISQSHGEG 133
 DB 384 LL--KVALDVAKGSMYLSHNNIIRDLKTANL-----MDHGLVYKVADEFVAVQIESG 436
 QY 134 ALGVE--GTPGQAPETPRIVYDEKVDMESTGNTVLLSGORPALGHOLQIA--KLLSK 191
 DB 437 VMTAETGYRMMAPEVEHKKADVFSTAYIWMELTSDIYAFLLTPLOAAVGVQK 496
 QY 192 GIRPVLGQPEEVEFRRLQALMECMTKPEKRPALSVSQM 234
 DB 497 GLRPKI--PKKTH--PKVGLLRCMHODPEORPLFEETIELMQ 536

RESULT 6
 J00229
 mixed-11neage protease kinase 1 - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
 C:Accession: S32467; J00229
 R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kreseer, T.
 Eur. J. Biochem. 213, 701-710, 1993
 A:Title: Identification of a new family of human epithelial protein kinases containi
 A:Reference number: S32467; MUID:93238756; PMID:8477742
 A:Accession: S32467
 A:Molecule type: mRNA
 A:Residues: 1-394 <DO2>
 C:Genetics:
 A:Gene: GDB:MLK1
 A:Cross-references: GDB:141921; OMIM:600136
 A:Map position: 14q24.3-14q31
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase ho
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific prote
 F:1-269/Domain: catalytic <CAT>
 F:1-268/Domain: catalytic <CAT>
 F:9-17/Region: protein kinase ATP-binding motif
 F:289-310/Region: leucine zipper motif
 F:324-345/Region: leucine zipper motif
 F:354-368/Region: basic

Query Match
 Best Local Similarity 7.28; Score 275.5; DB 2; Length 394;
 Matches 78; Conservative 42; Mismatches 92; Indels 37; Gaps 8;
 QY 3 RHLRATDAMKNFSEFQASMLHALQHCYIALGISIHP--LCFALPLPLSSINTVLS 60
 DB 33 RHDPDEDISQTIENVROEAKFLFAMLKHPIIALRGVCLKEPNLCIVMEFAGGPINRVLS 92
 QY 61 ENADSSFTPLGHMLTKI---AVQIASGLAYLHK---NIIFCDLKSNDILWMSL--- 110
 DB 93 -----GKRIPDILVMAVQIARGMYLHDEIVPILHRLKSSNLIQKVEN 141
 QY 111 -DVKEHINIKLSDGISQSHGEGALGVEGTPGYQAPETPRIVYDEKVDMESTGNTVLL 169
 DB 142 GDLNKKI-LKTTDGLAREHRTTKMSAAGTAYAMAPETPRIVYDEKVDMESTGNTVLL 200
 QY 170 LLSGQRALGHOLQIA-----KLLSKGIRPVLGQPEEVEFRRLQALMECMTKPEKRP 224
 DB 201 LLTGEVPRGIDGLAVGVAMNKLALPIPTCPERP-----FAKLMECCWNPDPHSP 253
 QY 225 LALSVSQM 233
 DB 254 SFTNIDQL 262

RESULT 7
 T04683
 hypothetical protein F8020.290 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
 C:Accession: T04683
 R:Bevan, M.; Rose, M.; Hempel, S.; Entlan, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Me

submitted to the Protein Sequence Database, July 1998

A:Reference number: 215381

A:Accession: T04683

A:Molecule type: DNA

A:Residues: 1-553 <BEV>

A:Cross-references: EMBL:AL031135

A:Experimental source: cultivar Columbia; BAC clone F8D20

C:Genetics:

A:Map position: 4

A:Introns: 69/2; 107/3; 176/2; 194/3; 231/3; 289/2; 325/2; 350/1; 408/3; 440/3; 467/3; 4

A:Note: F8D20.290

Query Match 7.1%; Score 270.5; DB 2; Length 553;

Best Local Similarity 30.3%; Pred. No. 1.5e-08;

Matches 74; Conservative 44; Mismatches 109; Indels 17; Gaps 8;

16 EFROASMLHLOHPCIALVIGISIH~PLCFALPLASSLNTVLSNARDSSPIPLGH 73

316 EFSDEVYIMRKVKNVQFAGCTRSPLCTVTEPMTRGSIYDPLKHN~KGVFKIQS 372

74 MLOKIAVOIASGLAYLHKKNIFCDLKSNDILVMSLDVKEHINIKLSDYGISRQSFHEG 133

373 LL~KVALDVSKMNYLHNNIHRDLTKANL~MDEHYVKAADFVAVRQTESG 425

134 ALGYE~GTPGQAPRIRIYDEKVDKMSGMVLYELSGRPAIGHQDIA~KKLSK 191

426 VMTMETGTWRMAPEVIMHKKPYDHRADVFSAIVLMEITGELPYSLYPLQAAVGVQK 485

192 GIRPVLGQPEEYQFRRLQALMECDTKREKRLALSVYSQKDPFATFMEELCCGQT 251

486 GIRPVI~PKETH~PKLTLELKKCKMOODPALRPFATTEMLNQLIREYIDLSLHKDKHG 542

252 AFFS 255

543 GYFS 546

RESULT 8

112956

hypothetical protein T6H20.50 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 22-Oct-1999

C:Accession: T12956

R:Choline, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Sa

submitted to the Protein Sequence Database, July 1999

A:Reference number: 217586

A:Accession: T12956

A:Molecule type: DNA

A:Residues: 1-1171 <CHO>

A:Cross-references: EMBL:AL096859; GSPDB:GN00061; ATSP:T6H20.50

A:Experimental source: cultivar Columbia; BAC clone T6H20

C:Genetics:

A:Gene: ATSP:T6H20.50

A:Map position: 3

A:Introns: 780/1; 857/3; 881/3; 932/3; 987/2; 1036/3; 1077/3; 1107/1

Query Match 7.1%; Score 270.5; DB 2; Length 1171;

Best Local Similarity 32.7%; Pred. No. 3.7e-08;

Matches 80; Conservative 42; Mismatches 100; Indels 23; Gaps 9;

6 BATDMKRFSEERQASMLHLOHPCIALVIGISIH~PLCFALPLASSLNTVLS 61

925 KPSDEERIDDFWNEAQNLAGLHPNVAFYGVVDSFGSVATVETVMVNSGLRNAOK 984

62 NARDSSPIPLGHMLTOKIAYOIASGLAYLHKKNIFCDLKSNDILVMSLDVKEHINIKLS 121

985 NVRN~PDRCKRQLAMQIAFMELHKGKTIYHFDLKSNDILVNLNLRDPIRP1~CAVG 1038

122 DYGISR~QSFHEGALVEGTPGQAPRI~RPRIVYDEKVDKMSGMVLYELSGORP 176

1039 DLGSLKVCQTLISG~GVRGLPMMAPDELNGTSLVSEKVDVFSFGVILMELPTGEPP 1096

177 ~ALGHQDIAKKLSKIRPVLGQPEEYQFRRLQALMECDTKREKRLALSVYSQKMD 235

1097 YADLHYGALLIGIVSNLIPQIPDFCDMDK~LLMERCSAEPSESPTEIYNELR~ 1152

236 PTFAT 240

1153 ~TMAI 1156

RESULT 9

C86273

protein kinases homolog F7A19.9 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001

C:Accession: C86273

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Al

Chid, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Ki

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marz

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C86273

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-438 <STO>

A:Cross-references: GB:AE005172; MID:95080776; PID:AD39286.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 7.0%; Score 268.5; DB 2; Length 438;

Best Local Similarity 31.5%; Pred. No. 1.5e-08;

Matches 74; Conservative 36; Mismatches 90; Indels 33; Gaps 8;

16 EFROASMLHLOHPCIALVIGISIH~PLCFALPLASSLNTVLSNARDSSPIPLGH 73

204 DFRHVDLVLKRNHNIVQFAGAVTERKPLMIFELVGLDGHQKRG~GLTP 257

74 MLOKIAVOIASGLAYLHKKN~IFCDLKSNDILVMSLDVKEHINIKLSDYGISR~ 127

258 TTAVFALDIAKGMVYLNENPVIIHRDLKERNVLLVSSAD~HLKVGFGSLIKY 314

128 ~QSFHE~GALVEGTPGQAPRIYDEKVDKMSGMVLYELSGORPALGHQDIA 185

315 QNSHDVYKMTGETGSYRMADEVRKRRYDKVDVFSFAMILYELMEGEPPANHEPEA 374

186 AKLSKIRPVLGQPEEYQFR~LQALMECDTKREKRLALSVYSQKMD 234

375 AKHVSQDGRPT~FRSKGCTPLRELIVKCMQADNQRSPFLDLKRL 421

RESULT 10

T12955

probable protein kinase (EC 2.7.1.-) T6H20.40 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jan-2000

C:Accession: T12955

R:Choline, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F

submitted to the Protein Sequence Database, July 1999

A:Reference number: 217586

A:Accession: T12955

A:Molecule type: DNA

A:Residues: 1-475 <CHO>

A:Cross-references: EMBL:AL096859; GSPDB:GN00061; ATSP:T6H20.40

A:Experimental source: cultivar Columbia; BAC clone T6H20

C:Genetics:

A:Gene: ATSP:T6H20.40

A:Map position: 3

A:Introns: 39/1; 226/3; 376/3

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase h

C:Keywords: phosphotransferase; protein kinase
F:159-422/Domain: protein kinase homology <KIN>

Query Match 7.0%; Score 268; DB 2; Length 475;

Best Local Similarity 29.3%; Pred. No. 1.8e-08;

Matches 66; Conservative 58; Mismatches 79; Indels 22; Gaps 9;

16 EFROASMLHLQHPICVALIGISHPICFALPLSLVSLMTVSENRDSEFIPLGML 75
DB EPIYATILSRSLSPNVKVGWNTGN-CITTEYPRSSLSYHLKLRCK--LPLEQIT 264
208 EPIYATILSRSLSPNVKVGWNTGN-CITTEYPRSSLSYHLKLRCK--LPLEQIT 264
QY 76 TOKIAYIASGLAYLHKNNIFCDLKSNDILWSDVHEHNINKLSDYISROSEFEGAL 135
DB 265 --DFGLDIAGKMEYIHSREIYHQDLKPEVNL-----INDFHLKADFGICEEYCDVL 317
QY 136 GVE-GTPGYOAPRIPIRYDEKVDMSYGMVYLLSGORALGHOLQIKKLS----- 190
DB 318 GDNITGYMMAPEVYKRRIPHGKCDYSGFGLLMEMVAG--ALPEYEEKFAEQIAYAVI 374
QY 191 -KGIRPVLGOPPEVQFRRLQALMECMPTKPEKRLALSVSOMK 234
DB 375 YKKIRPVY--PTDCP-AAKKELIERCWSQIDKREPMOYKAVLE 416

RESULT 11

F86297

hypothetical protein F309.7 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: F86297

R:Thelologists: A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzialis,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F86297

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 11147 <STO>

A:Cross-references: GB:AEO05172; NID:g4966348; PIDN:AD34679.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 7.0%; Score 267.5; DB 2; Length 1147;

Best Local Similarity 33.3%; Pred. No. 5.5e-08;

Matches 82; Conservative 41; Mismatches 88; Indels 35; Gaps 10;

QY 6 RATDANKNFESEFOEASMLHALQHPICVALIGISHPICFALPLSLVSLMTVSENRD 65
DB 900 RSEDERLTFGEWGEAEIISKLHPVAVFYGV-----VKDGGALAIY-FTYMD 950
QY 66 SFFPLGMLTOK-----IAYQIASGLAYLHKNNIFCDLKSNDILWSDVHEHNINKLSDYISROSEFEGAL 114
DB 951 GS---LRHVLVRRKRLDLRRKRLIAMDAPFGEYLAHANNIYHFDLCKDLNLVNLKDPGR 1007
QY 115 HINIKLSYGVIS---RQSHFEGALGVEGTPGYOAPRI--RPRIVYDEKVDMSYGMVYLLSGORALGHOLQIKKLS----- 169
DB 1008 PI-CKVGDGLSKIKRNTLVSG--GVKGTLPMAAPPELLNGSSSKYSEKVDVSEFGIVLWE 1064
QY 170 LLSGGRP-ALGHNQLQIAKLSKIGIRPVLGOPPEVQFRRLQALMECMPTKPEKRLALSVSOMK 228
DB 1065 ILTGEPPYANMHTGALITGIVNTLPTIPSTCDSDMR---ILMECCMAPNPAPRSFTE 1121
QY 229 VVSOMK 234
DB 1122 IAGRLR 1127

RESULT 12

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: H86179

R:Thelologists: A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzialis,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H86179

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 11028 <STO>

A:Cross-references: GB:AEO05172; NID:g2494111; PIDN:AB80620.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 6.9%; Score 263; DB 2; Length 1029;

Best Local Similarity 30.6%; Pred. No. 8.8e-08;

Matches 78; Conservative 46; Mismatches 89; Indels 42; Gaps 10;

QY 7 ATDAKKNFESEFOEASMLHALQHPICVALIGISHPICFALPLSLVSLMTVSENRD 66
DB 768 SSEQARQKDFREARILANLHHPVAVFYGV-----YVDGGGTATVTEYVNGS 839
QY 67 SFFPLGMLTOK-----IAYQIASGLAYLHKNNIFCDLKSNDILWSDVHEHNINKLSDYISROSEFEGAL 119
DB 840 ---LRHVLVRRKRLDLRRKRLIAMDAPFGEYLAHANNIYHFDLCKDLNLVNLKDPGR 894
QY 120 LSDYIS---RQSHFEGALGVEGTPGYOAPRI--RPRIVYDEKVDMSYGMVYLLSGORALGHOLQIKKLS----- 172
DB 895 VDFGLSKIKRNTLVSG--GVKGTLPMAAPPELLNGSSSKYSEKVDVSEFGIVLWE 950
QY 173 GORPALGHR-----OLQIAKLSKIGIRPVLGOPPEVQFRRLQALMECMPTKPEKRLALSVSOMK 224
DB 951 GEPPYANMHTGALITGIVNTLPTIPSTCDSDMR---ILMECCMAPNPAPRSFTE 1121
QY 225 IALSVSOMK 239
DB 1008 SFTFIERLSMTVA 1022

RESULT 13

protein-tyrosine kinase (EC 2.7.1.112) - slime mold (Dictyostelium discoideum)

C:Species: Dictyostelium discoideum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Feb-2000

C:Accession: T18287

R:Adler, K.
submitted to the EMBL Data Library, July 1996

A:Reference number: T18287

A:Accession: T18287

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 11338 <ADL>

A:Cross-references: EMBL:U64830; NID:g1468982; PID:g1468983; PIDN:AB804999.1

C:Genetics:

A:Insertions: 1181/3

C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match 6.9%; Score 263; DB 2; Length 1338;

Best Local Similarity 28.6%; Pred. No. 1.2e-07;

Matches 65; Conservative 51; Mismatches 97; Indels 14; Gaps 4;

QY 9 DAMNFESEFROEASMLALQHPICVIALIGISIHPLCFALAPLASSLNTVLSNARDSS 68
 DB 1089 ESIISNDKREIKESVSLKSHHPNVYTEMGARIDPPCIFTETLQAGSYDVHLIOK----- 1143
 QY 69 IPIGHMLTOKIAVOIASGLAVLHKNNIIFCDLKSNDILWMSLDVKEHINIKLSYGISRO 128
 DB 1144 IKLNPILMYMKIHDLSIGMEHLSTQMLHRDLTSKNIL-----VDERKNIKIAOFGIATY 1198
 QY 129 SFHEGALGVETPGYQAPRIPIVDEKYMFSYGNVLYELLSGQRPALG-HHQLQIAK 187
 DB 1199 LSDDMTSLGITNPRMRSPELTKGLVYNEKVDVYSFGLVYVEIYGIKIPFEDLDTASAK 1258
 QY 188 KLSKGRIPVVGPEEVOFRRLQALMECMTKPKRPLAIVSOMK 234
 DB 1259 AAFENRPAIPDCPVSILRK---LITCMASDPSQRPSEITLLELE 1302

RESULT 14

A53800
 mixed-lineage protein kinase (EC 2.7.1.1) 3 - human
 N/Alternate names: protein kinase PTK1; protein kinase SPRK
 C/Species: Homo sapiens (man)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: A53800; 158395
 R/Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
 J. Biol. Chem. 269, 15092-15100, 1994
 A/Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont
 A/Reference number: A53800; M0ID:94253068; PMID:8195146
 A/Accession: A53800
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-847 <GALD>
 A/Cross-references: GB:U07747; NID:9464027; PIDN:AA19647.1; PID:9464028
 R/ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Laszlam, N.J.
 Oncogene 9, 1745-1750, 1994
 A/Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domai
 A/Reference number: 158395; M0ID:94259754; PMID:8183572
 A/Accession: 158395
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-847 <RES>
 A/Cross-references: GB:L2976; NID:9488295; PIDN:AA59859.1; PID:9488296
 C/Genetics:
 A/Genes: GDB:MLK3; PTK1; SPRK
 A/Cross-references: GDB:134755; OMIM:600050
 A/Map position: 11q13.1-11q13.3
 C/Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
 C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
 F:48-100/Domain: SH3 homology <SH32>
 F:115-383/Domain: protein kinase homology <KIN>
 F:123-131/Region: protein kinase ATP-binding motif
 F:403-424/Region: leucine zipper motif
 F:438-459/Region: leucine zipper motif
 F:468-482/Region: basic

Query Match 6.8%; Score 260.5; DB 1; Length 847;
 Best Local Similarity 22.6%; Pred. No. 9.8e-08;
 Matches 162; Conservative 84; Mismatches 223; Indels 249; Gaps 33;

QY 18 ROEASMLALQHPICVIALIGISIHPLCFALAPLASSLNTVLSNARDSSPILGHML 75
 DB 162 ROEASMLALQHPICVIALIGISIHPLCFALAPLASSLNTVLSNARDSSPILGHML 215
 QY 76 TOKIAVOIASGLAVLHKNNIIFCDLKSNDILWMSLDVKEHINIKLSYGISRO 129
 DB 216 V-NNVAOIAAGMYLHCALVPIYHRDLKSNILLLQPSIDDMKHTLKITTOFGLAREM 274
 QY 130 FHGAGLVETPGYQAPRIPIVDEKYMFSYGNVLYELLSGQRPALGHHQLQIAK 186
 DB 275 HKTQSSAAGTYAMAPPEVIAKSTFSGSDVSWFGLVYVEIYGIKIPFEDLDTASAK 334
 QY 187 --KLSKGRIPVVGPEEVOFRRLQALMECMTKPKRPLAIVSOMKDPFAFMTE 244

DB 335 AVNKLTLPISTCEP-----FAQLMADCMADPHRRRDFASIILOL-EALEAOVLRE 386
 QY 245 ICCGKQAFAPSOGQEVTVFWMD-KESRNYTVNTEKGLMEVORCCGMVSOLOV 303
 DB 387 M---PRDSFSMO-----BEWKREI-----QGLFELR---AAEKELLR 420
 QY 304 QRSIMTATEDOKIYIYTLKGNCPLNFPQALDPVAVYCELAIVIKNSYLVLAGIAD 363
 DB 421 EEEELTRARARQ-----SOA-----EQLRREHL---LAQV 448
 QY 364 LVAVF-----PVT---RGPKDCSYLCSHTANKSFIADE----- 397
 DB 449 ELEVEFERETLLQOVDREPHVRRRGTYK-----RSKLARADGGERISMP 496
 QY 398 -----DANONYPVK-----AMEVNSGSVWVSNPGLLYID 430
 DB 497 DFKRITVQASPGIDRRNRNFEVGPDSPTFPFRPAIQLEPAEG-CAWGROSP----- 549
 QY 431 CASLEICRLEPYMAPSMVTSVVCSSG-----RGEVYV 465
 DB 550 -----RLED-----SSNGERACMAMPSSPKRGEAONGRRRMDERTW 590
 QY 466 CUDKANSLVYISTTYQLCARFCGYPF-LDMFPVPLDTEPPAASHTANPKYPEGD 524
 DB 591 YLSDSDSPGSPSTPALNG---NPFRSLPEEPKRPVPAERSSSGT--PKLIORA 644
 QY 525 SIADVSIMYSEEGTOILHQESLTDYCSMSYSSSPRROAAPS-----LPSP 576
 DB 645 LMGTLALASLGIGRL-----QPGGPRGRGSGPTTPPTPAAPCTEP 680
 QY 577 ASSGSVPSTDCSDMLNTP-----GAASDRSEHD-----LTPMDGETFS 617
 DB 691 PPSPLICFSIKTDPSPPTAPLLDLGIPGVSASKSPREERPGTVSPRGTSRS 748

RESULT 15

B96827
 hypothetical protein TBK14.1 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C/Accession: B96827
 R/Theologis, A.; Ecker, J.R.; Palm, C.T.; Federspiel, N.A.; Kaul, S.; White, O.; Al
 Chalm, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
 ansen, N.F.; Hughes, B.; Hulzar, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kl
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marz
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tai
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; M0ID:21016719; PMID:11130712
 A/Accession: B96827
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1248 <STO>
 A/Cross-references: GB:AE005173; NID:94835752; PIDN:AA030219.1; GSPDB:GN00141
 C/Genetics:
 A/Genes: TBK14.1
 A/Map position: 1

Query Match 6.8%; Score 260.5; DB 2; Length 1248;
 Best Local Similarity 32.9%; Pred. No. 1.6e-07;
 Matches 81; Conservative 41; Mismatches 89; Indels 35; Gaps 10;

QY 6 RATDAKNSERROEASMLALQHPICVIALIGISIHPLCFALAPLASSLNTVLSNARD 65
 DB 1001 RSEDERLTGEFGAEILSKLHPVAVFYGV-----VKDGGGLAVY-TEYMD 1051
 QY 66 SSPFPGHMLTK-----IAYOLASGLAVLHKNNIIFCDLKSNDILWMSLDVKE 114
 DB 1052 GS---LRHVLVKRDRHLDRKRLLIIMDAAFGEYVLSHNTVFDLKCNDLNLKDPGR 1108

Tue Apr 15 14:19:11 2003

us-09-836-392-21.rpr

Page 7

```

Qy 115 HINKLSDYGIS---RQSFHEGALGVGEKPGVQAPFI--RPIIVDEKVDNFSGMVLVE 169
Db 1109 PI-CRVADFGSLKSKRNTLVSG--GVRGTLPMNAPPELLNGSSSKVSEYVDVDFSGIYLME 1165
Qy 170 LLSGCRP-ALGHHOLQIAKTLKSGIRPRVGPQEEVOFRILALMECDWTRPEKRPILALS 228
Db 1166 ILTDEEYPAANHHGAILIGGIYVNTLTLPITIPGCDDEMR---TLMECNAPNPMARPSYTE 1222
Qy 229 VVSQMK 234
Db 1223 IAGRLR 1228

```

Search completed: April 8, 2003, 10:26:26
Job time : 62 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:13:58 ; Search time 27 Seconds
(without alignments)
1118.325 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822
Sequence: 1 MNRHLPATDMKMFSEPRDE.....IFQSTEEELGRLACTRRKR 728

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	285	7.5	954	M3KA_HUMAN	Q02779 homo sapien
2	275.5	7.2	394	M3K9_HUMAN	P80192 homo sapien
3	257	6.7	410	KYK2_DICDI	P18161 dictyostell
4	254.5	6.7	506	SRK4_SPOLA	P26680 spongillia 1
5	246	6.4	937	KORL_MOUSE	Q92138 mus musculu
6	243.5	6.4	505	FRK_HUMAN	P42683 mus musculu
7	238.5	6.2	412	APKB_ARATH	P46573 arabidopsis
8	238.5	6.2	1307	PHYL_CERPU	P25848 ceratodon p
9	238	6.2	937	RORL_HUMAN	Q01973 homo sapien
10	236.5	6.2	497	SPK1_DUGTI	P42687 dugesia tig
11	236	6.2	821	CTRL_ARATH	Q05609 arabidopsis
12	235.5	6.2	323	FLK_RAT	P09760 rattus norv
13	235	6.1	505	SRK1_SPOLA	P42686 spongillia 1
14	231.5	6.1	805	FUSE_DROME	P23647 drosophila
15	231.5	6.1	822	FER_HUMAN	P16591 homo sapien
16	228.5	6.0	579	M3K7_MOUSE	Q62073 mus musculu
17	228.5	6.0	606	M3K7_HUMAN	Q43318 homo sapien
18	228	6.0	859	M3K3_HUMAN	Q12852 homo sapien
19	228	6.0	888	M3K3_MOUSE	Q60700 mus musculu
20	226	5.9	888	M3K3_MOUSE	Q63796 rattus norv
21	225.5	5.9	410	APKA_ARATH	Q06548 arabidopsis
22	225	5.9	968	SR10_HUMAN	Q94804 homo sapien
23	224.5	5.9	943	ROR2_HUMAN	Q01974 homo sapien
24	223.5	5.8	1584	KYK1_DICDI	P18160 dictyostell
25	222	5.8	507	ICK_CHICK	P42683 gallus galli
26	222	5.8	552	SRK1_DROME	P00528 drosophila
27	221.5	5.8	389	NAK_ARATH	P43293 arabidopsis
28	221.5	5.8	628	KSYK_PIG	Q06655 sus scrofa
29	219.5	5.7	635	KSYK_HUMAN	P43405 homo sapien
30	218.5	5.7	629	KSYK_MOUSE	P48025 mus musculu
31	218.5	5.7	1477	KYK2_MOUSE	Q25197 hydra atten
32	218	5.7	2347	KROS_HUMAN	P08922 homo sapien
33	217	5.7	714	HUNK_MOUSE	O88866 mus musculu

ALIGNMENTS

```

RESULT 1
ID M3KA_HUMAN STANDARD; PRT; 954 AA.
AC Q02779; Q12761; Q14871;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 10 (EC 2.7.1.37)
DE (Mixed lineage kinase 2) (Protein kinase MST).
GN MAP3K10 OR MKK2 OR MST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96128179; PubMed=8536694;
RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
RA Sutherland G.R., Simpson R.J.;
RT "Complete nucleotide sequence, expression, and chromosomal
RT localisation of human mixed-lineage kinase 2."
RL Eur. J. Biochem. 234:492-500(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95249256; PubMed=7731697;
RA Katoh M., Hirai M., Sugimura T., Terada M.;
RT "Cloning and characterization of MST, a novel (putative)
RT serine/threonine kinase with SH3 domain."
RL Oncogene 10:1447-1451(1995).
RN [3]
RP SEQUENCE OF 244-480 FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=9328756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains."
RL Eur. J. Biochem. 213:701-710(1993).
RN [4]
RP CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
RC TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
RN [5]
RP SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
RN [6]
RP MAP KINASE KINASE SUBFAMILY.
RN [7]
RP SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X90846; CA62351.1;
CC EMBL: Z48615; CA88531.1;
CC PIR: S32468; S32468.

```


DR InterPro: IPR001245; Tyr-kinase.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00069; kinase; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk-kinase; 1.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 1.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 1.
 DR Transferrase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
 SH3 domain; Phosphorylation; Alternative splicing.
 FT DOMAIN 54 116
 FT DOMAIN 122 214
 FT NP_BIND 240 493 PROTEIN_KINASE.
 FT BINDING 246 254 ATP (BY SIMILARITY).
 FT BINDING 268 268 ATP (BY SIMILARITY).
 FT ACT_SITE 359 359 BY SIMILARITY.
 SQ SEQUENCE 506 AA; 57561 MW; 8ABDF4A2546C280B CRC64;

Query Match 6.7%; Score 254.5; DB 1; Length 506;
 Best Local Similarity 30.3%; Pred. No. 1e-10; Matches 73; conservative 46; Mismatches 87; Indels 35; Gaps 8;

OY 13 NSEFRQENSMALQHPICALIGISIT--HPLCEALIELAPISLNTLVSENARDSSFLP 70
 DB 276 SIEEFLEERSIKQKQHPHLLQLVAVCTKEEPIYVTELMKGLSLEYLRGGRSLKLPD 335
 OY 71 LGHMLTKIAYQIAGSLAYLHKNNIIFOLKSDNIIWGLDKVKEHNINLSYGISR--- 127
 DB 336 LVDKMS---OVASGMSTLEQONTHRLAARNIL---VGENHICVADFGARVID 385
 OY 128 OSFHEGALVEGTPCYQAPETIRPIYDEKVFMSGYVLVELLS-GORPALGHHQLQIA 186
 DB 386 EETIYAKIGAKFKPIKMTAPEAMYSRFTIKSDVWSGIVLYEVITYGRTYGGMTRNAQVL 445
 OY 187 KLLSGIR--PVLGQPEYQFRRLQALMECCDPTKPEKPLALSVYSQMKDPPTATEMTE 244
 DB 446 EQLQOSTYRMPDPCPE---KLVAIMDCWRDPASR-----PTFFELSMQ 488
 OY 245 L 245
 DB 489 L 489

RESULT 5
 ROR1_MOUSE STANDARD; PRT; 937 AA.

AC Q92139; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
 DE (EC 2.7.1.113) (Neurotrophic tyrosine kinase, receptor-related 1)
 DE (MORF).
 GN ROR1 OR NTRKRL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A. Pubmed=10233192;
 RA Oishi T., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,
 RA Hata T., Akira S., Matsuda Y., Yamamura H., Otsu H., Minami Y.;
 RT "Spatio-temporally regulated expression of receptor tyrosine kinases,

RT mRor1, mRor2, during mouse development: implications in development
 RT and function of the nervous system.";
 RL Genes Cells 4:41-56(1999).
 CC -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHOSE ROLE IS NOT YET
 CC CLEAR.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 KIRINLE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB010383; BAA75480.1; -
 CC HSSP: P00747; ICEA.
 DR MGI: MGI:1347520; Ror1.
 DR InterPro: IPR000719; Euk-kinase.
 DR InterPro: IPR000024; Fz-domain.
 DR InterPro: IPR003066; Ig-MHC.
 DR InterPro: IPR003598; Ig-C2.
 DR InterPro: IPR000001; kirinle.
 DR InterPro: IPR001245; Tyr-kinase.
 DR Pfam: PF00047; Ig; 1.
 DR Pfam: PF00051; kirinle; 1.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF01392; Fz; 1.
 DR PRINTS: PR00018; KIRINLE.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk-kinase; 1.
 DR ProDom: PD000395; Kirinle; 1.
 DR SMART: SM00408; IgC2; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS0038; Fz; 1.
 DR PROSITE: PS00021; KIRINLE_1; FALSE_NEG.
 DR PROSITE: PS50070; KIRINLE_2; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR Transferrase; Tyrosine-protein kinase; ATP-binding; Receptor;
 KW transmembrane; Signal; Glycoprotein; Kirinle; Phosphorylation;
 KW Immunoglobulin domain.
 FT CHAIN 1 29
 FT SIGNAL 1 29
 FT CHAIN 30 937
 FT DOMAIN 30 406
 FT TRANSMEM 407 427
 FT DOMAIN 428 937
 FT DOMAIN 73 139
 FT DOMAIN 165 299
 FT DOMAIN 312 391
 FT DOMAIN 473 746
 FT DOMAIN 753 782
 FT DOMAIN 784 851
 FT NP_BIND 853 876
 FT BINDING 479 487
 FT BINDING 506 506
 FT ACT_SITE 615 615
 FT MOD_RES 645 645
 FT DISULFID 79 131
 FT CARBOHYD 47 47
 FT CARBOHYD 66 66
 FT CARBOHYD 184 184
 FT CARBOHYD 315 315
 FT POTENTIAL.
 FT TYROSINE-PROTEIN KINASE TRANSMEMBRANE
 FT RECEPTOR ROR1.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE C2-TYPE DOMAIN.
 FT FZ.
 FT KIRINLE.
 FT PROTEIN KINASE.
 FT SER/THR-RICH.
 FT SER/THR-RICH.
 FT ATP (BY SIMILARITY).
 FT ATP (BY SIMILARITY).
 FT BY SIMILARITY.
 FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT BY SIMILARITY.
 FT N-LINKED (GLCNAc) (POTENTIAL).
 FT N-LINKED (GLCNAc) (POTENTIAL).
 FT N-LINKED (GLCNAc) (POTENTIAL).
 FT N-LINKED (GLCNAc) (POTENTIAL).
 FT N-LINKED (GLCNAc) (POTENTIAL).

SQ SEQUENCE 937 AA; 104156 MW; D728733E67D1782C CRC64;
 Query Match 6.4%; Score 246; DB 1; Length 937;
 Best Local Similarity 25.3%; Pred. No. 9.8e-10;
 Matches 79; Conservative 54; Mismatches 103; Indels 76; Gaps 10;

QY 2 LRHRRDAMKNESEFQESMHALQPCIVALLIG--ISTHPLC-----FA 46
 DB 505 IKTLDKYNPPQOWTEFQOEASLMELHHPNIVCLGAVTQGPVCMLEFYNGQDLHEFL 564
 QY 47 LELAPLSLNTVLSENARDSFFIFLGHMTRKRIAYQIASGLAYLHKRIIFCDKSDNLL 106
 DB 565 IMRPHSDVGSSDEPDGVKSLDHGDEL--HIMQIAGKEVYSNHFVKDLAARILL 622
 QY 107 VMSLDVKEHNIKLSVDIGISFQSGHEGALGVEGTGPGYQAFIRPRIV---YDEKVDMS 162
 DB 623 -----IGQLHVKISDLGLSREIYSADYRVSQSSKSLPIRMMPPEALMYGKSSDSIDMS 677
 QY 163 YGMVLVELLS--GORPALGHQLOJAKKLKSGIRPVLOGPPEVQRRLOALMMEGMDTFKE 221
 DB 678 FGVLVMEIFSGLOPYGFSNQEVIEMVRK--KOLLPSGDCP--PRMYSLMECMEIIPS 734
 QY 222 RRP-----LALSVYSOMKDPFATFMYELC 246
 DB 735 RRPFKDILHVRKMEGLSHTSSTPSGNGATQTSLSASPVSNLSNPRFPNFM----- 790
 QY 247 CGKQTAFFSSOG 258
 DB 791 -----FPSOG 795

RESULT 6
 FRK_HUMAN STANDARD; PRT; 505 AA.
 ID FRK_HUMAN P4685; Q13128;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE TYROSINE-PROTEIN KINASE FRK (EC 2.7.1.112) (Nuclear tyrosine protein kinase FRK).
 GN FRK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Lymphoid; PubMed-7510261;
 RX MEDLINE-94171047; PubMed-7510261;
 RA Lee J., Wang Z., Luoh S.-M., Wood W.I., Scadden D.T.;
 RT "Cloning of FRK, a novel human intracellular SRC-like tyrosine kinase-encoding gene.";
 RL Gene 138:247-251(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95210168; PubMed-7966183;
 RA Cance W.G., Craven R.J., Bergman M., Xu L.H., Alltalo K., Liu E.T.;
 RT "Rak, a novel nuclear tyrosine kinase expressed in epithelial cells.";
 RL Cell Growth Differ. 5:1347-1355(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Urinary bladder;
 RA Strausberg R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE-95293373; PubMed-8099900;
 RA Cance W.G., Craven R.J., Weiner T.M., Liu E.T.;
 RT "Novel protein kinases expressed in human breast cancer.";
 RL Int. J. Cancer 54:571-577(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- TISSUE SPECIFICITY: RESTRICTED TO CELLS LINES DERIVED FROM TISSUES OF LYMPHOID, BRAIN, BREAST, COLON AND BLADDER ORIGIN.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DB EMBL; U00803; AAA18284.1; .
 DB EMBL; U22322; AAC50116.1; .
 DB EMBL; BC012916; AAH12916.1; .
 DB HSSP; P00523; 2PTK.
 DB Genew; HGNC:3955; FRK.
 DB KIM; 606573; .
 DR InterPro; IPR000719; Euk_Pkinase.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR001245; Tyr_Pkinase.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_Pkinase; 1.
 DR ProDom; PD000066; SH3; 1.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS50001; SH2; 1.
 DR PROSITE; PS50002; SH3; 1.
 DR Transfaser; Tyrosine-protein kinase; ATP-binding; SH2 domain; SH3 domain; Phosphorylation.
 KW SH3 domain; Phosphorylation.
 FT DOMAIN 42 110
 FT DOMAIN 116 208
 FT DOMAIN 234 491
 FT NP_BIND 240 248
 FT BINDING 262 262
 FT ACT_SITE 354 354
 FT MOD_RES 387 387
 FT VARIANT 122 122
 FT /FTID=VAR_006783.
 FT G->R.
 FT CONFLICT 115 115 P->A (IN REF. 2).
 SQ SEQUENCE 505 AA; 56254 MW; 06EC05DDBCD930B CRC64;
 Query Match 6.4%; Score 243.5; DB 1; Length 505;
 Best Local Similarity 30.5%; Pred. No. 6.3e-10;
 Matches 75; Conservative 49; Mismatches 77; Indels 45; Gaps 11;

QY 15 SEPRQASMLHALQPCIVALLIGIST--HPLCFLELAPLSLNTVLSENARDSFFPLG 72
 DB 272 NDFREAOIMKLNHHPKLIQIYAVCTLEDPIYITELMRHSLQBYL-QNDGSKI----- 326
 QY 73 HMLQKTI--AQIASGLAYLHKRIIFCDKSDNLLVMSLDVKEHNIKLSVDGISR--- 127
 DB 327 -HLVQVDMAAQVAVSNAVYIESRNYIHDDLAARVL-----VGEHNIYKVADEFLARVK 380
 QY 128 -----OSFHEGALGVEGTGPGYQAFIRPRIVDEKVMFSYGMVLVELLS--GORPALG 179
 DB 381 VDNEDYIESRHEITLPLVYKT-----APEAIRSNKRSIKSDVWSPFILLLEITTYGKMYSG 436
 QY 180 HHQIQAIRKISGIRPVLOGPPEVQRRLOALMMEGMDTFKEKRRPLALSVYSOMKDPTEA 239

DB 437 MTGADVTCMLAQNFR--LPQSPNCP-QCFYINIMLECMNAEPKRR-----PTEF 481
 QY 240 TENYEL 245
 DB 482 TLRWKL 487

RESULT 7

APKB_ARATH STANDARD; PRT; 412 AA.
 ID APKB_ARATH STANDARD; PRT; 412 AA.
 AC P46573; Q9SLH5; (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein kinase APK18 (EC 2.7.1.-).
 GN APKB OR AT2G28930 OR P194.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Beutlo M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.V., Cressy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana";
 RT Nature 402:761-768(1999).
 RL [2]
 RN SEQUENCE OF 143-346 FROM N.A.
 RP STRAIN=cv. Columbia;
 RX MEDLINE=93081726; PubMed=1450380;
 RA Hirayama T., Oka A.;
 RA "Novel protein kinase of Arabidopsis thaliana (APK1) that
 phosphorylates tyrosine, serine and threonine";
 RT Plant Mol. Biol. 20:653-662(1992).
 RL -1- FUNCTION: POSSIBLE BI-FUNCTIONAL KINASE. IN VITRO, IT EXHIBITS
 SERINE/THREONINE ACTIVITY. IN VIVO, CAN PHOSPHORYLATE TYROSINE
 CC RESIDUES OF LIMITED SUBSTRATES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AC005315; AAC3221.1;
 DR EMBL: D10152; BA20968.1;
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; TYR_pkinase.
 DR Pfam: PF00069; pkinase.1
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: P000001; Euk_pkinase.1.
 DR SMART: SM00221; STYK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
 KW ATP-binding; Multigene family; Myristate.

FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
 FT DOMAIN 69 356 PROTEIN KINASE.
 FT BIND 75 83 ATP (BY SIMILARITY).
 FT BINDING 107 107 ATP (BY SIMILARITY).
 FT ACT_SITE 204 204 BY SIMILARITY.
 SQ SEQUENCE 412 AA; 45746 MW; EB1CA0B1A626A5DA CRC64;

Query Match 6.2%; Score 238.5; DB 1; Length 412;
 Best Local Similarity 32.5%; Pred. No. 1.1e-09;
 Matches 83; Conservative 37; Mismatches 92; Indels 43; Gaps 11;

QY 9 DAMKNSEFEQROASMLHALQHPICVALLIGISI---HPLCFALAPLPLSLNTVLEENARD 65
 DB 112 DGMQGHQEWLAVNVNLGQESHNLVLYLCYDEDEHRL-LYEFNPRGSLNHLRR--RG 168
 QY 66 SSFPLGMLTKQIKVQIASGLAYLH--KKNIFDGLSDNLTWVSLDKHEINIKLSDY 123
 DB 169 SYFOPLSWTLRLKVALGAKGIAFLHMETSVIYDFKTSNILLDS---EYNAKLSDF 223
 QY 124 GISROSFHEGALG-----VEGTPGYOAPIRIRIVYDEVDMESYGMVLYELLSGOR 175
 DB 224 GLAK---DGPFGDSDSHVSTRIMGTGYAPEYLAATGHTTSDVYSIGVVLLEVLGSR 279
 QY 176 PALGHQQLQIAKLSKGRIPVLGQPEYQFR---RLQ-----ALMECWDYK 219
 DB 280 -AVDKNRPPGEOKIVEMARPLLNKRKL-FVYIDNRLQOYGMERACKVATLALCLTFE 337
 QY 220 PEKRPALSYVSQMK 234
 DB 338 ITRPMEVSHLE 352

RESULT 8

PHY1_CERPU STANDARD; PRT; 1307 AA.
 ID PHY1_CERPU STANDARD; PRT; 1307 AA.
 AC P25848; P93100; (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Light-sensor Protein kinase [includes: Phytochrome; Protein kinase
 (EC 2.7.1.-)].
 GN PHY1 OR PHX.
 OS Ceratodon purpureus (Moss).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 OC Bryopsida; Dicranidae; Dicranales; Dittichaceae; Ceratodon.
 CC NCBI_TaxID=3225;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93099252; PubMed=1463836;
 RA Thummmler F., Dutner M., Kreis P., Dittlich P.;
 RA "Molecular cloning of a novel phytochrome gene of the moss Ceratodon
 purpureus which encodes a putative light-regulated protein kinase";
 RL Plant Mol. Biol. 20:1003-1017(1992).
 RN [2]
 RP SEQUENCE OF 49-538 FROM N.A.
 RX MEDLINE=9108543; PubMed=2261981;
 RA Thummmler F., Beetz A., Ruediger W.;
 RA "Phytochrome gene in the moss Ceratodon purpureus using the polymerase
 RT chain reaction";
 RT FEBS Lett. 275:125-129(1990).
 RN [3]
 RP REVISIONS TO C-TERMINUS.
 RC STRAIN=WT3;
 RA Pasentis K., Paulo N., Dittlich P., Algarra P., Thummmler F.,
 RA Dutner M., Kreis P.;
 RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: REGULATOR PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT
 CC ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS
 CC MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT
 CC ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN
 CC PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS
 CC RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE

```

OY 67 SEIPGHLTKIVQVLSGLYVLAHKNNIFCDKSDNLIWMSLDVKE-----HINIKSD 122
Db 1095 DSDPEISILEVVDIIQTSSEGMNLYLHEKEGIIRHDLKSNAILKSVAYTSEISGYVAVKAD 1154
OY 123 YGISR-----QSFHEGA-----LGVGGTGGYQAPDIRIRIYDEKV----- 158
Db 1155 FGLSKTNSQSTRYSNQTFNNRGTHNRMADEVINLGYESTEG-----EISDGKVRKYP 1206
OY 159 ---DMESTGYMYLLEISGCRPALGHHQQLAKKL-SKGIKPLVGGPEVQFRLLQALME 214
Db 1207 LKSDYSSFGMAYCEVLTGDVPEPEEKNNNNNKRWYLEGVRPDLPHACPIE---LKALITD 1263
OY 215 CWDPRKKRPLALSVSQMKDPTAFATMYELCCGKQTAFFSSQGGQGEY 261
Db 1264 CWNDDPLKRPSPFAVICQKLK-----YLYLMTKG-----FSSYQDSY 1300

RESULT 9
RORL_HUMAN STANDARD; PRT; 937 AA.
AC Q01973; Q92776;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
   (EC 2.7.1.11) (Neurotrophic tyrosine kinase, receptor-related 1).
GN ROR1 OR NTRK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RX MEDLINE=93100347; PubMed=13344954;
RA Maslakowski P., Carroll R.D.;
RT "A novel family of cell surface receptors with tyrosine kinase-like
   domain.";
RL J. Biol. Chem. 267:26181-26190(1992).
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RX MEDLINE=97030043; PubMed=8875995;
RA Reddy U.R., Phatak S., Pleasure D.;
RT "Human neural tissues express a truncated Ror1 receptor tyrosine
   kinase, lacking both extracellular and transmembrane domains.";
RL Oncogene 13:1555-1559(1996).
CC -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHOSE ROLE IS NOT YET
   CLEAR.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
   tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A LONG FORM (SHOWN HERE) AND A
   SHORT FORM/T-ROR1. ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED STRONGLY IN HUMAN HEART, LUNG, AND
   KIDNEY, BUT WEAKLY IN THE CNS. THE SHORT ISOFORM IS STRONGLY
   EXPRESSED IN FETAL AND ADULT CNS AND IN A VARIETY OF HUMAN
   CANCERS, INCLUDING THOSE ORIGINATING FROM CNS OR PNS
   NEUROECTODERM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING EARLY
   EMBRYONIC DEVELOPMENT. THE EXPRESSION LEVELS DROP STRONGLY AROUND
   DAY 16 AND THERE ARE ONLY VERY LOW LEVELS IN ADULT TISSUES.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
   SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).

```



```

DR EMBL: M97675; AAA60275.1;
DR EMBL: U38894; AAC50714.1;
DR HSP: P00747; ICEA.
DR Genew: HGNC:10256; ROR1.
DR MIM: 602336;
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR000024; Fz_domain.
DR InterPro: IPR00306; Ig_MHC.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF01392; Fz; 1.
DR PRINTS: PRO0018; KRINGLE.
DR PRINTS: PRO0109; TYRKINASE.
DR PRODOM: PD000001; Euk_Pkinase; 1.
DR PRODOM: PD000395; Kringle; 1.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00219; TYRC; 1.
DR PROSITE: PS50038; Fz; 1.
DR PROSITE: PS50021; KRINGLE_1; FALSE_NEG.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50109; PROTEIN_KINASE_TYR; 1.
DR Transfaser: Tyrosine-protein kinase: ATP-binding; Receptor;
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW Immunoglobulin domain;
FT SIGNAL 1 29
FT CHAIN 30 937
FT
FT DOMAIN 30 406
FT TRANSSEM 407 427
FT DOMAIN 428 937
FT DOMAIN 73 139
FT DOMAIN 165 299
FT DOMAIN 312 391
FT DOMAIN 473 746
FT DOMAIN 753 782
FT DOMAIN 784 851
FT DOMAIN 853 876
FT NP_BIND 479 487
FT BINDING 506 506
FT ACT_SITE 615 615
FT MOD_RES 645 645
FT DISULFID 79 131
FT CARBOHYD 47 47
FT CARBOHYD 66 66
FT CARBOHYD 184 184
FT CARBOHYD 315 315
FT VARSPIC 1 549
SO SEQUENCE 937 AA; 104312 MW; 0D0694BDF29F4773 CRC64;

Query Match 6.28; Score 238; DB 1; Length 937;
Best local Similarity 25.38; Pred. No. 3.6e-09;
Matches 79; Conservative 53; Mismatches 104; Indels 76; Gaps 10;

```

```

OY 163 YGAVLYELLS-GORPALGHOLJAKLSKIRPVLCOPPEVORRLOALMEGMDRKE 221
DB 678 FGVLMEIFSRGLQPYGFGFSNOEVIEMVR--ROLPCSEDCP-PRVYSLMEGMDRIPS 734
OY 222 KRP-----LALSVYOKMDPFAFATFYEIC 246
DB 735 RRPFRKDIHVALRSMEGLSSTSTSPSGNATTQTTSLASPSVSNLSNPRPYM----- 790
OY 247 CGKQTAFSSQSG 258
DB 791 -----FPSQSG 795

RESULT 10
SPK1_DUGT1 STANDARD; PRT; 497 AA.
ID SPK1_DUGT1
AC P42687;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase Spk-1 (EC 2.7.1.112).
OS Dugesia tigrina (Planarian).
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Sertata; Tricladida;
OC Paludicola; Dugesidae; Girardia.
OX NCBI_TaxID=6162;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94181283; PubMed=7510865;
RA Bugaya F., Garcia-Fernandez J., Rulfort M., Baguna J., Salo E.;
RT "Structure and expression of Spk-1, an src-related gene product found
RT in the planarian Dugesia (G) tigrina."
RL Oncogene 9:1267-1272(1994).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC
CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL: X75310; CA53058.1;
DR HSP: P00523; 2PDK.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00069; Pkinase; 1.
DR PRODOM: PD000001; Euk_Pkinase; 1.
DR PRODOM: PD000066; SH3; 1.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00326; SH3; 1.
DR SMART: SM00219; TYRC; 1.
DR PROSITE: PS50107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
KW Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding;
KW SH2 domain; SH3 domain.
FT DOMAIN 33 94
FT DOMAIN 100 200
FT DOMAIN 220 482
FT NP_BIND 226 234
ATP (BY SIMILARITY).

```

```

FT BINDING 248 248 ATP (BY SIMILARITY)
FT ACT_SITE 342 342 D -> A.
FT VARIANT 212 212 H -> D.
FT VARIANT 235 235
SQ SEQUENCE 497 AA; 57447 MW; 412FD1058697524 CRC64;

Query Match 6.2%; Score 236.5; DB 1; Length 497;
Best Local Similarity 27.8%; Pred. No. 1.9e-09;
Matches 66; Conservative 52; Mismatches 84; Indels 35; Gaps 9;

QY 16 EFROEASMLHALOHPICVALIGISIH---LCPALFELAPLSLNTVLSN-ARDSSFIIP 70
DB 259 KFLTEAKIMKRLHRKRLVLELGVCTEPQDMPMLIYEMKNGSKEXELKPKDPCKTLNQ 318
QY 71 LGMHLQKIAVOIASGLAYLHKRIKIFCDKSDNLIWMSLDVKEHINIKISDYGISQSF 130
DB 319 MVHMA-----EISEGNALIESEKRYVRDRADNLT-----VANDLRKVADELTELT- 367
QY 131 HEGALGVEGTP-----GYOAPETPRIVYDEKVMESYGMVLYELLS-GORPALGHQL 183
DB 368 -DPSLGGQEKTLRFPRKMTAPAKSKFTSVDWYSGIVMEILLTMASSPPIPAK 426
QY 184 QIAKLSKGRIPVLGOPEEYQF-----RLQALMECDTPEKPKPLASVWSQ 233
DB 427 EVIEKYSKGYR---MENPEKFTIGVCCPDEIKYIMTWCDANPEKRPTELVLOEK 479

RESULT 11.
CTRL_APATH STANDARD; PRT; 821 AA.
ID CTRL_APATH
AC 005609;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase CTRL (EC 2.7.1.37).
GN CTRL OR AT5G03730 OR F17C15.150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. Columbia; TISSUE=Seedling;
RX MEDLINE=93161417; PubMed=8431946;
RA Kobata M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,
RA Huang A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Stuberhan K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stuberhan K., Peplin K., Speich J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney M., Dante M.,
RA Du H., Edwards J., Fryman J., Haakenen B., Imar E., Lattrell P.,
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Woliam A., Yoakum M., Bell M., Dedila N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Yil D., Baker J.,
RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
RA Martensen R., McComble M.R., Wilson R.K., Murphy G., Bancroft I.,
RA Voickert G., Hambutt R., Duesterhoef A., Stekema W., Pohl T.,
RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grynopriest B., Zimmermann W.,
RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirkse W., Moolman P., Klein Lankhorst R.,

```

```

RA Weltzenegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gleen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rüd S., Schoof H.,
RA Scheller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
RT *Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT thaliana.
RL Nature 408:823-826 (2000).
CC -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
CC PATHWAY.
CC -1- CATALYTIC ACTIVITY: ATP + a protein - ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
CC -1- MISCELLANEOUS: CTRL MOTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
CC COTYLEDON GROWTH IS IMPAIRED.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MTL/RAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L08789; AAA32779.1; -
DR EMBL; L08790; AAA32780.1; -
DR EMBL; AL162506; CAB82838.1; -
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR004040; STY_Pkinase.
DR InterPro; IPR002290; Ser_thr_Pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR SMART; SM00221; STYKC1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR KMW; Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 65 69 POLY-GLY.
FT DOMAIN 135 141 POLY-GLY.
FT DOMAIN 531 809 PROTEIN KINASE.
FT NP_BIND 557 565 ATP (BY SIMILARITY).
FT BINDING 578 578 ATP (BY SIMILARITY).
FT ACT_SITE 676 676 BY SIMILARITY.
FT MUTAGEN 596 596 E->K: IN CTRL-4; EXHIBITS ETHYLENE-
FT TREATED PHENOTYPE.
FT MUTAGEN 694 694 D->E: IN CTRL-1; EXHIBITS ETHYLENE-
FT TREATED PHENOTYPE.
SQ SEQUENCE 821 AA; 90306 MW; 2922D3DCDC15BC CRC64;

Query Match 6.2%; Score 236; DB 1; Length 821;
Best Local Similarity 29.0%; Pred. No. 4.2e-09;
Matches 72; Conservative 45; Mismatches 75; Indels 56; Gaps 10;

QY 15 SEFROEASMLHALOHPICVALIGISIH---LCPALFELAPLSLNTVLSN-ARDSSFIIP 71
DB 591 NEFLREVAIKRLHRKRLVLELGVCTEPQDMPMLIYEMKNGSKEXELKPKDPCKTLNQ 318
QY 72 GMLTQKIAVOIASGLAYLHKRIKIFCDKSDNLIWMSLDVKEHINIKISDYGISR-- 127
DB 647 DERRRLMAYDVAAGMYLHNRNPIYVRDLKSPNLT-----YDKKTVKVCDFGLSRK 701
QY 128 QSFHGAISGVEGTPGYOAPETPRIVYDEKVMESYGMVLYELLSGSR-- 176
DB 702 ASYLLSKSAAGTEWMAPEVLRLDEPSNEDSVSGVILWELATLLOPWCNLPAPVOYA 761
QY 177 ALGH--HOLQIAKLSKGRIPVLGOPEEYQF---RLQALMECDTPEKPKPLASVWSQ 234
DB 762 AVGKCKRLEIRPLNLF-----GVAAIIIEGWTNPEPKR----- 795
QY 235 DPTATFV 242
DB 796 -PSFATIM 802

```

RESULT 12

```

FLK_RAT      STANDARD;          PRT;          323 AA.
ID   FLK_RAT
AC   P09760;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Tyrosine-protein kinase FLK (EC 2.7.1.112) (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Mistar; TISSUE=Brain;
RX   MEDLINE=94167102; PubMed=248525;
RA   Letvin K., Yee S.P., Pawson T.;
RT   "Novel protein-tyrosine kinase cDNAs related to fps/fes and eph
RT   cloned using anti-phosphotyrosine antibody."
RL   Oncogene 3:621-627(1988).
CC   -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC   tyrosine phosphate.
CC   -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. FES/FPS
CC   SUBFAMILY.
CC   -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X13412; CAA31778.1;
DR   PIR; S04328; S04328.
DR   HSSP; P11362; IFK.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001452; SH3.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   ProDom; PD000093; SH2; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50001; SH2; PARTIAL.
DR   PROSITE; PS50002; SH3; PARTIAL.
KW   Transferase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;
KW   SH2 domain.
FT   NON_TER      1
FT   DOMAIN       <1 51
FT   NP_BIND      64 315
FT   NP_BIND      70 78
FT   NP_BIND      92 92
FT   ACT_SITE     185 185
FT   MOD_RES      215 215
SQ   SEQUENCE 323 AA; 37104 MW; D7BABDE50C3EAC1 CRC64;

```

Query Match

Best Local Similarity 32.6%; Score 235.5; DB 1; Length 323;
Matches 72; Conservativity 36; Mismatches 82; Indels 31; Gaps 9;

```

QY 16 EFROEASMLHALQHPICVALIGISI--HPLCFALDELAPLSLTVLSENAKRSFPLGH 73
DB 105 KFLQEKILKQYDHPNIVKLIGCTGROPVYIIMELVPGGDFLKRKRDLEKLQQL--162
QY 74 MLTGKIAVQIASGLAVLHKNNIFCDLKSNDNIIVMSLDVKEHINIKLSIDGISRSQSFHSG 133

```

```

DB 163 ---VFESLDVAAAGMLYLEGKNCIHRDLAARNCL-----VGENNTLTISDFGSKR--EDG 212
QY 134 ALGVECTPG-----YQAEIRPRIVYDEKYMESYGVLYELLS-GORPALGHQLOI 185
DB 213 --GVYSSGLKQPIKWTAPALNVRYSSESQVMSFGILLMETFSLGCPYFGMTNQCA 270
QY 186 AKLSKGRIPVLAQ--PEVQFRRLQALMECDWTKPERP 224
DB 271 REQVERGYRMSAPQNCPEEI-----FTIMKCMWDYKPERNP 306

```

RESULT 13

```

SRK1_SPOLA      STANDARD;          PRT;          505 AA.
ID   SRK1_SPOLA
AC   P42686;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Tyrosine-protein kinase SRK1 (EC 2.7.1.112).
GN   SRK1.
OS   Spongilla lacustris (Freshwater sponge).
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
OC   Haplosclerida; Spongillidae; Spongilla.
OX   NCBI_TaxID=6055;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92334872; PubMed=1378585;
RA   Oltjie S., Raulf F., Barnekow A., Hannig G., Scharlt M.;
RT   "Multiple src-related kinase genes, srk1'4, in the fresh water sponge
RT   Spongilla lacustris."
RL   Oncogene 7:1625-1630(1992).
CC   -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC   tyrosine phosphate.
CC   -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -1- ALTERNATIVE PRODUCTS: 2 isoforms; SRK1 and SRK4 (AC P42680); are
CC   produced by alternative splicing.
CC   -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
CC   SUBFAMILY.
CC   -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X61601; CAA43798.1;
DR   HSSP; P00523; 2PK.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001452; SH3.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF00018; SH3; 1.
DR   ProDom; PD000093; SH2; 1.
DR   SMART; SM00252; SH2; 1.
DR   SMART; SM00326; SH3; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50001; SH2; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;

```

KW SH3 domain; phosphorylation; alternative splicing.
 FT DOMAIN 54 116
 FT DOMAIN 122 214
 FT DOMAIN 240 493
 FT NP_BIND 246 254
 FT BINDING 248 268
 FT ACT_SITE 359 359
 FT SEQUENCE 505 AA; 57693 MW; 63999E411C6112P91 CRC64;
 Query Match 6.1%; Score 235; DB 1; Length 505;
 Best Local Similarity 28.1%; Pred. No. 2,5e-09;
 Matches 70; Conservative 52; Mismatches 91; Indels 36; Gaps 10;
 QY 16 EFGQESMLHALQPCIVALLIGISI--HPICFALPLASLTMTVSENNARSSFFILGH 73
 DB 279 EFGQESIMKRLRPKLIQILYACVTEEPYITELMKYGLS--LEYLNGEGVLTKEQ 335
 QY 74 MLOKIAIYQIASGLAYLHKNNIFCDLKSNDILWISLWVKEHINIKLSDYISR--QSF 130
 DB 336 LV--DVAAQVAGSMSTLEQOQNYIHRDLAANNIL--VEHGICVADRGARVIDEEL 388
 QY 131 HEGALVEGTPGYQAPETIRRYIDKVDKMSYGVLYELLS--GORPALGHOLQIAKL 189
 DB 389 YEHTAKPKPIKMTAPEAAMYNRTIKSDVSGVLYEITTYGRRPYPMPTPEVLEKI 448
 QY 190 SKGIRPVLAGPEEVQFRRLQALMECHMDTKERKPLALSVSQMKDPTATMYELCCGK 249
 DB 449 QQVTR--MPCPANC-P-KQFDHMDCKREDPASR-----PTFETLQML----- 489
 QY 250 QTAFFSSQG 258
 DB 490 -EEFFNSEG 497
 RESULT 14
 ID FUSE_DROME STANDARD; FRT: 805 AA.
 AC P23647; Q9VWR7; Q27412; Q26346; Q26347;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase fused (EC 2.7.1.-).
 GN FU OR CG6551.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=Oregon-R; TISSUE=Ovary, Embryo, and Pupae;
 RX MEDLINE=94206843; PubMed=8155575;
 RA Thermond P., Bussion D., Guillemet E., Limbourg-Bouchon B., Preat T.,
 RA Terracol R., Tricoire H., Lamour-Ignard C.;
 RT "Molecular organization and expression pattern of the segment polarity
 gene fused of Drosophila melanogaster.";
 RL Mech. Dev. 44:65-80(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceuliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotcher P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Moadary C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacible J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
 RA Palenbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski R.P., Smith T.,
 RA Splier E., Spreading A.C., Stapleton M., Strongi R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wastarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 1-155 FROM N.A., AND MUTAGENESIS.
 RX MEDLINE=94140077; PubMed=8307322;
 RA Preat T., Therond P., Limbourg-Bouchon B., Pham A., Tricoire H.,
 RA Bussion D., Lamour-Ignard C.;
 RT "Segmental polarity in Drosophila melanogaster: genetic dissection of
 fused in a Suppressor of fused background reveals interaction with
 costal-2.";
 RL Genetics 135:1047-1062(1993).
 RN [4]
 RP SEQUENCE OF 1-273 FROM N.A., AND FUNCTION.
 RX MEDLINE=90370097; PubMed=2168522;
 RA Preat T., Therond P., Lamour-Ignard C., Limbourg-Bouchon B.,
 RA Tricoire H., Erk I., Mariol M.-C., Bussion D.;
 RT "A putative serine/threonine protein kinase encoded by the segment-
 polarity fused gene of Drosophila.";
 RL Nature 347:87-89(1990).
 CC -1- FUNCTION: PROBABLE SERINE/THREONINE-PROTEIN KINASE; MATERALLY
 CC REQUIRED FOR CORRECT PATTERNING IN THE POSTERIOR PART OF EACH
 CC EMBRYONIC MEMBER. MAY BE INVOLVED IN THE CONTROL OF CELL DIVISION
 CC DURING METAMORPHOSIS AND OVARIAN DEVELOPMENT. MAY INTERACT WITH
 CC COSTAL-2.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL IMAGINAL DISKS, HIGHER LEVEL
 CC IN WING DISK.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY,
 CC LOW EXPRESSION IS PRESENT IN MALES, LARVAE AND PUPAE.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X60468; CAA56640.1; -
 CC EMBL: L34782; AA28552.1; -
 CC EMBL: AE003509; AAF48671.1; -
 CC EMBL: S69165; AAB29840.1; -
 CC EMBL: S69166; AAB29841.1; -
 CC EMBL: X55759; CAA39285.1; -
 CC PIR: S11380; S11380.
 CC HSSP: Q00534; 1B18.
 CC FBASE: F89N001079; fu.
 CC InterPro: IPR000719; Euk_pkinase.
 CC InterPro: IPR002290; Ser_thr_pkinase.

Tue Apr 15 14:19:10 2003

us-09-836-392-21.rapb

Page 8

OY 247 CGKOTAFPSOG 258
Db 791 -----FPSOG 795

Search completed: April 8, 2003, 10:34:42
Job time : 50 secs

Tue Apr 15 14:19:10 2003

us-09-836-392-21.rapb

Pa.

Db 437 MGAQVYQIMLAQVYR--LTPQSNCP--QOFYNIMLECNAPKBER-----PTFE 481
QY 240 TFWYEL 245
Db 482 TLRWKL 487

RESULT 13

US-09-757-982-5
; Sequence 5, Application US/09757982
; Patent No. US20020094539A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: INT-050
; CURRENT APPLICATION NUMBER: US/09757,982
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-982-5

Query Match 6.4%; Score 243; DB 10; Length 455;
Best Local Similarity 27.9%; Pred. No. 3.1e-11;
Matches 69; Conservative 52; Mismatches 100; Indels 26; Gaps 7;

QY 10 AMNFESEFROEASMLHALQHPCLVALIGISHPICFAL--ELAPLSLNTVLSNARDSS 67
Db 43 AVKLLKEAEELSLVSHRNITQFYGVLEPNYGVYFASLSLYDYNR--SE 100
QY 68 FPLGHLTKRIAYQIASGLAYLHKR--NIIFCDLSDNIIWVSDVKEHINIKLSDYG 124
Db 101 EMDMDHIMTW--ATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADV----LKIDCFG 153
QY 125 ISROSFEAGLVGEGTGYQAPETPRIVYDEKYMFSYGVLYELLSGORPALGHOLQ 184
Db 154 ASREHNTHTMSLVGTTPMNAPEVIOQLPVSECTDYSYGVLEMLTREVFPKGLBGLQ 213
QY 185 IAKLSGIRPVLGQPEVQF-----RRLOALMECWDTRKPRPLALSVYSOKKDPFA 239
Db 214 VAWL-----VVEKNERLTIIPSSCPSPFAELHQCWEADAKKRPFRQIISLESMSND 266
QY 240 TFWYELC 246
Db 267 TSLPDKC 273

RESULT 14

US-09-904-389-2
; Sequence 2, Application US/09904389
; Patent No. US20020129404A1
; GENERAL INFORMATION:
; APPLICANT: Clendennen, Stephanie K.
; TITLE OF INVENTION: CTRI HOMOLOGUE FROM MELON
; FILE REFERENCE: 4257-0029.30
; CURRENT APPLICATION NUMBER: US/09/904,389
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/218,307
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Cucumis melo
; FEATURE:

NAME/KEY: VARIANT
; LOCATION: (154)...(154)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-09-904-389-2

Query Match 6.3%; Score 242.5; DB 10; Length 850;
Best Local Similarity 30.0%; Pred. No. 8.2e-11;
Matches 69; Conservative 49; Mismatches 93; Indels 19; Gaps 8;

QY 15 SEFROEASMLHALQHPCLVALIGISHP--LCPALAPLSLNTVLSNARDSSFPILG 72
Db 617 NEFLREVALIMKSRHPVIVLEMGAVTRPVLSTVETLSGSLRLHK-----SGVKDID 672
QY 73 HMLTKRIAYQIASGLAYLHKR--NIIFCDLSDNIIWVSDVKEHINIKLSDYGISRSP 130
Db 673 ETRINMAFPAVAKGMNLIHRDPPVIRDLKSPYL-----VKKYVVKCDGSLSLKA 727
QY 131 HE--GALVGEPTGYQAPETPRIVYDEKYMFSYGVLYELLSGORPALGHOLQAKK 188
Db 728 RFLSKSKAGTEPMNAPEVLRDEPSNEKSDVYSFGVILWELATLQOPMCNLNPAQVYAA 787
QY 189 LS-KGIRPVLGQPEVQFRRLOALMECWDTRKPRPLALSVYSOKKDPF 237
Db 788 VGEKGR--LDIPRDVN--PKLASLVACMADEPKRPSFSIMETTLKPMT 834

RESULT 15

US-09-974-298-129
; Sequence 129, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huel-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 129
; LENGTH: 937
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1331526CD1
US-09-974-298-129

Query Match 6.2%; Score 238; DB 9; Length 937;
Best Local Similarity 25.3%; Pred. No. 2.1e-10;
Matches 79; Conservative 53; Mismatches 104; Indels 76; Gaps 10;

QY 2 LRHRATDAMKNESEFROEASMLHALQHPCLVALIG--ISHPIC-----FA 46
Db 505 IKTLKDYNNPQOMMEFQDEASLMAELHNPVLCIGAVTOBPVCMLEFYINGDLHEFL 564
QY 47 LELAPLSLNTVLSNARDSSFPILGHLTKRIAYQIASGLAYLHKRNIIFCDLSDNII 106
Db 565 IMRSPHSDVCGSSDSDGVKSSLDHGDPL--HIALQIAGHEYSNHFVAKDLAARIL 622
QY 107 VMSLDVKEHINIKLSDYGISROSFEAGLVGEGTP---GYQAPETPRIVYDERVDMFS 162
Db 623 -----IGBQLHVAKISDGLSREITSADYRVQSKSLPIRMMPPEALMYGKFSDDSDIWS 677
QY 163 YGVNLYELLS--GQRPALGHOLQIAKLSKIRPVLPQPEVQFRRLOALMECWDTRPE 221
Db 678 FGVLWELFEFSGLOPYGFSNQEVIEWVR--ROLPCSEDCP--PRYSIMLTGCMNEIPLS 734
QY 222 KRP-----LALSVYSOKKDPFAFPMYELC 246
Db 735 RRPKRDIVRLRSWEGISHTSTTPSGGNATTOTTSLSASPVNSLNPRIPNYM----- 790

OTHER INFORMATION: kinase 3
US-09-977-261-6

Query Match
Best Local Similarity 30.5%; Score 243.5; DB 9; Length 505;
Matches 75; Conservative 49; Mismatches 77; Indels 45; Gaps 11;

15 SEFROEASMLALHPCVVALIGISI--HPLCFALPLSLMTVSENRDSEFIPLG 72
DB 272 NDFLEAQMIMKLRHKLQIYAVCTLEDPIYITELMRHSLQRYL-QNDTGSKI----- 326
QY 73 HMLTOKI--AVOIASGLAYLHKNNIIFCDLSNDILVSLDVKEHINIKLSDYGISR--- 127
DB 327 -HLTQOVMAQVAGSMAYLESRYIHRLDLAARNVL-----VGEHNIYKVADEFLARYEK 380
QY 128 -----SFHEGALGVEGTGPGYQAPETRPRIYDEKYMESYGMVLYELLS-GGRPALG 179
DB 361 VDNEIYRSREIKLPVMT-----APEAIRSKNSFKSDVMSFGILLIETTYGKMPYSG 436
QY 180 HHQLOIAKLSKGRIPVLGPEEVOFRRLQALMECDTKPEKRPALSVSOMKDPTEFA 239
DB 437 MTGAQVIOMLAQNYR--LPQPSNCP-QQFYINIMLECNNAEPRER-----PTFE 481
QY 240 TFMVEL 245
DB 482 TLRWKL 487

RESULT 11
US-09-977-269-6
Sequence 6, Application US/09977269
Patent No. US20020082037A1
GENERAL INFORMATION:
APPLICANT: ULARICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SORES, IRMINGARD
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977, 269
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232, 545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 505
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
OTHER INFORMATION: kinase 3
US-09-977-269-6

Query Match
Best Local Similarity 30.5%; Score 243.5; DB 10; Length 505;
Matches 75; Conservative 49; Mismatches 77; Indels 45; Gaps 11;

15 SEFROEASMLALHPCVVALIGISI--HPLCFALPLSLMTVSENRDSEFIPLG 72
DB 272 NDFLEAQMIMKLRHKLQIYAVCTLEDPIYITELMRHSLQRYL-QNDTGSKI----- 326
QY 73 HMLTOKI--AVOIASGLAYLHKNNIIFCDLSNDILVSLDVKEHINIKLSDYGISR--- 127
DB 327 -HLTQOVMAQVAGSMAYLESRYIHRLDLAARNVL-----VGEHNIYKVADEFLARYEK 380
QY 128 -----SFHEGALGVEGTGPGYQAPETRPRIYDEKYMESYGMVLYELLS-GGRPALG 179
DB 361 VDNEIYRSREIKLPVMT-----APEAIRSKNSFKSDVMSFGILLIETTYGKMPYSG 436
QY 180 HHQLOIAKLSKGRIPVLGPEEVOFRRLQALMECDTKPEKRPALSVSOMKDPTEFA 239
DB 437 MTGAQVIOMLAQNYR--LPQPSNCP-QQFYINIMLECNNAEPRER-----PTFE 481

QY 240 TFMVEL 245
DB 482 TLRWKL 487

RESULT 12
US-09-982-610-20
Sequence 20, Application US/09982610
Patent No. US20020146420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-982-610-20

Query Match
Best Local Similarity 30.5%; Score 243.5; DB 10; Length 505;
Matches 75; Conservative 49; Mismatches 77; Indels 45; Gaps 11;

15 SEFROEASMLALHPCVVALIGISI--HPLCFALPLSLMTVSENRDSEFIPLG 72
DB 272 NDFLEAQMIMKLRHKLQIYAVCTLEDPIYITELMRHSLQRYL-QNDTGSKI----- 326
QY 73 HMLTOKI--AVOIASGLAYLHKNNIIFCDLSNDILVSLDVKEHINIKLSDYGISR--- 127
DB 327 -HLTQOVMAQVAGSMAYLESRYIHRLDLAARNVL-----VGEHNIYKVADEFLARYEK 380
QY 128 -----SFHEGALGVEGTGPGYQAPETRPRIYDEKYMESYGMVLYELLS-GGRPALG 179
DB 361 VDNEIYRSREIKLPVMT-----APEAIRSKNSFKSDVMSFGILLIETTYGKMPYSG 436
QY 180 HHQLOIAKLSKGRIPVLGPEEVOFRRLQALMECDTKPEKRPALSVSOMKDPTEFA 239

RESULT 8
US-10-143-133-2
; Sequence 2, Application US/10143133
; Publication No. US20020197658A1
; GENERAL INFORMATION:
; APPLICANT: Joganathan, Thillainathan
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use
; FILE REFERENCE: KINE-023
; CURRENT APPLICATION NUMBER: US/10/143,133
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,555
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ. ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 2
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-143-133-2

Query Match 6.8%; Score 260.5; DB 9; Length 847;
Best Local Similarity 22.6%; Pred. No. 3e-12;
Matches 162; Conservative 84; Mismatches 223; Indels 249; Gaps 33;

QY 18 ROEASMLHALQHPICVALIGISIH--LCFALPLAPLSINTVLSENARDSFPLGLHML 75
DB 162 ROEASMLHALQHPICVALIGISIH--LCFALPLAPLSINTVLSENARDSFPLGLHML 75
QY 76 TORIAVOIASGLAVLHKK--NIIFCDLKSNDILVW--SLDYKEHINIKLSDYGISROS 129
DB 216 V--NMAVOIANGMHYHCEALVPIHDLKSNITLLQPISSDDEHNTKLTIDGLAREW 274
QY 130 FHGALGVBETPGIOAPEIRIVYDEKVDMEFYGMVLYELLSQAPALHHLQIA--- 186
DB 275 HKTQSSAGTYAMAPAEVITKASTFSGSDVMSFGVLTWELTGEVYRIGIDCLAVAYGV 334
QY 187 --KLSKGIKIPVLPGEPEVOFRRLQALMECMTPEKRPPLASVVSOMDPFATPME 244
DB 335 ANVKLLIPITSTCEP-----PAQLMDCMADPHRRRDFASITLQI--EALPAQVIRE 386
QY 245 LCCGKQATFESSQGEYTVFWDG--KEESRNYTVNTEKGLMEVORNCCEPMKVSQLOV 303
DB 387 M---PRDSFHSQC-----EGMKREI-----QGLFDELK-----AKKELLSR 420
QY 304 QBSLMTATEQKIYIITLKMCPLNTPQALDIPAVVTCFLAVFYIKKSYLVLAGADG 363
DB 421 EEEELTRAREROR-----SOA-----EQLRRREHL-----LAOW 448
QY 364 LVAVF-----PVY---RGTPKDSYLSHNTANSKFSIADG----- 397
DB 449 ELEVEFERELTLLOQVDREPRHVRRRRGTFR-----RSKLARDGGRIRISIMPL 496
QY 398 -----DARONDYPVK-----AMEVYVSGSEVWY SNGPGLIYID 430
DB 497 DFKHRTIVQASPGIDRRRRNVEVGPDSPTFPRFRAIQLPABRG--QAMGQOSP----- 549
QY 431 CASLEICRLEPYMAPSMTSVVCSSEG-----RGEVYVW 465
DB 550 -----RLED-----SSNGERRACWAMPSSPKPEAONGRRRRSMDRATW 590
QY 466 CLDDKANSLVWVHSTYQOLCARFYCGVSP--LRDMFVPRPLDTPPPASHATANKVEGD 524
DB 591 YLSDSSSPGSPSTPALNG-----NPPRSLEPEEKRPVPAERSSSGT--PKLQRA 644
QY 525 SIADVSIMYSEELGTQILLHQESITDYCSMSYSSSPRROAARSPSS-----LPSSP 576
DB 645 LIRGTALLASIGLRDL-----QPPGGGREGREGSEPTTPPTPACTPCTEP 690
QY 577 ASSSVFSTDCDSMDLHTP-----GAASDRSHND-----LIPMDETS 617
DB 691 PPSPLICFSLKTPDSPTPAFLDLIDGIPVQGRSAKSPRREERGGTVPSPPTSSRS 748

RESULT 9
US-09-977-260-6
; Sequence 6, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ. ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ. ID NO 6
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
US-09-977-260-6

Query Match 6.4%; Score 243.5; DB 9; Length 505;
Best Local Similarity 30.5%; Pred. No. 3.3e-11;
Matches 75; Conservative 49; Mismatches 77; Indels 45; Gaps 11;

QY 15 SEFROEASMLHALQHPICVALIGISI--HPLCFALPLAPLSINTVLSENARDSFPLGL 72
DB 272 NDFLRQIMKLNLRHPLILQVAVCTLEDPIIITTELMRGSIQEYI--QNDTSKI----- 326
QY 73 HMLTQKI--AYQIASGLAVLHKKNIIFCDLKSNDILVWSDYKEHINIKLSDYGISR-- 127
DB 327 -HITQCYDMAAOVAGMAYLESRYIHRDLAARVNL-----VEGHNTYKVADEGLAVFK 380
QY 128 -----QSFHEGALGVESGPGYQAPRIYVDEKVDMEFYGMVLYELLS--GORPALG 179
DB 381 VNDEDIYESHHEIKLPKWT-----APPAISNKFISIDVMSFGILYELIITTKGMYSG 436
QY 180 HHOLOIAKLSKGIKIPVLPGEPEVOFRRLQALMECMTPEKRPPLASVVSOMKDEPTFA 239
DB 437 MTGAQVIOQLAQNRR--LRPDSNCP--QQFYNIMLECMANAPKER-----PTFE 481
QY 240 TEMTEL 245
DB 482 TLRWKL 487

RESULT 10
US-09-977-261-6
; Sequence 6, Application US/09977261
; Publication No. US20030054527A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1259
; CURRENT APPLICATION NUMBER: US/09/977,261
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ. ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ. ID NO 6
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte

```

QY 61 ENARDSFFPLGMLTOKIAVOIASGLAYLHKNNIFCDLKSNDILWMSLDYKEHINIKL 120
DB 1413 ENARDSFFPLGMLTOKIAVOIASGLAYLHKNNIFCDLKSNDILWMSLDYKEHINIKL 1472
QY 121 SDYGISROSFHEGALGVEGTGPGYOAPETIRPRIYDEKVMFSYGMVLYELLSGORPALGH 180
DB 1473 SDYGISROSFHEGALGVEGTGPGYOAPETIRPRIYDEKVMFSYGMVLYELLSGORPALGH 1532
QY 181 HOLOIAKTSKGIKIRPVLGPEEVOFRRLQALMECHDTPERKRPALSLVYSQMKDPTFAT 240
DB 1533 HOLOIAKTSKGIKIRPVLGPEEVOFRRLQALMECHDTPERKRPALSLVYSQMKDPTFAT 1592
QY 241 FYEELCCGKQTAFFSSGOGYEYTVFWDGKESRNTYVNTTEKGLMEVOHMCPCGMKVSQ 300
DB 1593 FYEELCCGKQTAFFSSGOGYEYTVFWDGKESRNTYVNTTEKGLMEVOHMCPCGMKVSQ 1652
QY 301 LOVORSMTATDOKIYITTLKMGCPPLNTPOQALDTPAVVTCFLAVPVIKKNSYVLACL 360
DB 1653 LOVORSMTATDOKIYITTLKMGCPPLNTPOQALDTPAVVTCFLAVPVIKKNSYVLACL 1673
QY 361 ADGLVAFPVYVGTGPRDSCSYLCSHTANRSKFSIADEDARONPYPKAMEVYVNSGSEVY 420
DB 1674 ADGLVAFPVYVGTGPRDSCSYLCSHTANRSKFSIADEDARONPYPKAMEVYVNSGSEVY 1733
QY 421 SNGPGLVTDCASLIEICRLEPYMAPSMYTSYVCSSEGGEEVYVCLDOKANSIYVYHST 480
DB 1734 SNGPGLVTDCASLIEICRLEPYMAPSMYTSYVCSSEGGEEVYVCLDOKANSIYVYHST 1793
QY 481 TYOLCARFVCGVPSPLRDMFPVRPLDTEPPASHTANPYVPGSDIADVSIMYSEELGTQ 540
DB 1794 TYOLCARFVCGVPSPLRDMFPVRPLDTEPPASHTANPYVPGSDIADVSIMYSEELGTQ 1853
QY 541 ITHOESLTDYCSMSYSSSPROAARSPSLPSBPASSSVFSTDCSDMLTHPGAA 600
DB 1854 ITHOESLTDYCSMSYSSSPROAARSPSLPSBPASSSVFSTDCSDMLTHPGAA 1913
QY 601 SDRSEHDLTPMDGEFFSOHLQAVKILAVBDLIWPRRGDVIYIGLEKDEAGRGVIAV 660
DB 1914 SDRSEHDLTPMDGEFFSOHLQAVKILAVBDLIWPRRGDVIYIGLEKDEAGRGVIAV 1973
QY 661 LKARELTPGVYDAVAVAKOVVCFEENENTEMCLAVRGAGARFDIFYOSYEEGLYL 720
DB 1974 LKARELTPGVYDAVAVAKOVVCFEENENTEMCLAVRGAGARFDIFYOSYEEGLYL 2033
QY 721 EACTRR 727
DB 2034 EACTRR 2040

```

```

RESULT 6
US-09-862-027-19
; Sequence 19, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-19

```

```

Query Match 7.2%; Score 275.5; DB 10; Length 394;
Best Local Similarity 31.3%; Pred. No. 6,6e-14;
Matches 78; Conservative 42; Mismatches 92; Indels 37; Gaps 8;

```

```

QY 3 RHLRATDANKNSEFEROASMLHALOHCIALVIGISHP--LCFALSLAPLSLNTVLS 60
DB 33 RHDPDEDSITQENNRKOEKFLAMLRHPIILDRGCLMCLYMERARCGPLRRVLS 92
QY 61 ENARDSFFPLGMLTOKI---AVOIASGLAYLHKNNIFCDLKSNDILWMSLDYKEHINIKL 110
DB 93 -----GKRIPDILVMAVAVIARGMNYLHDEALVPIIHRDLKSNLILLOKVEN 141
QY 111 -DYKEHINKLSDYGISROSFHEGALGVEGTGPGYOAPETIRPRIYDEKVMFSYGMVLYE 169
DB 142 GDLSNRI-LKIDFGIAREMHTTKMSAAGTYAMAPPEVIRASMSKSGSDVSYGLWME 200
QY 170 ILSGORPALGHQLOIA-----KLSKGIKIRPVLGPEEVOFRRLQALMECHDTPERKRP 224
DB 201 LITGEVPRFGIDGLRAVAVANKLALPIPTSCPEP-----FAKLMECDNPPDHSRP 253
QY 225 IALSVYSQ 233
DB 254 SFTNILDQL 262

```

```

RESULT 7
US-10-014-882-2
; Sequence 2, Application US/10014882
; Patent No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Kleke, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encodin
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1036)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-10-014-882-2

```

```

Query Match 6.9%; Score 263; DB 12; Length 1036;
Best Local Similarity 32.4%; Pred. No. 2,6e-12;
Matches 81; Conservative 42; Mismatches 91; Indels 36; Gaps 10;

```

```

QY 9 DAMKNSEFEROASMLHALOHCIALVIGISHP--LCFALSLAPLSLNTVLS-SENA-- 63
DB 160 DAAAAESVRRARERFAMLRHPIILDRGCLMCLYMERARCGPLRRVLS 219
QY 64 -----RDSSTIPGLHMLTOKIAVOIASGLAYLHKNNIFCDLKSNDILWMSLDYKEHINIKL 113
DB 220 DRRAGPRARIRP-PHIVY-KMAVOIARGMNYLHDEALVPIIHRDLKSNLIL--LENI 275
QY 114 EHINTI-----KLSDYGISROSFHEGALGVEGTGPGYOAPETIRPRIYDEKVMFSYGMVLYE 168
DB 276 EHDIDCNKTLKITDGLANEMHRTTKMSAAGTYAMAPPEVIRASMSKSGSDIWSYGLVM 335
QY 165 ELSGORPALGHQLOIA-----KLSKGIKIRPVLGPEEVOFRRLQALMECHDTPERKRP 223
DB 336 ELITGEVPRFGIDGLRAVAVANKLALPIPTSCPEP-----FAKLMECDNPPDHSRP 253
QY 224 PLALSVYSQ 233
DB 389 PSFALILEDL 398

```

Tue Apr 15 14:19:10 2003

us-09-836-392-21.rab

QY 121 SDVIGISROSFHEGALGVEGPGYQAPETIRRIYDEKVDMSYGMVLYELLSCORPLGH 180
DB 1446 SDVIGISROSFHEGALGVEGPGYQAPETIRRIYDEKVDMSYGMVLYELLSCORPLGH 1505
QY 181 HOLOIAKLSKSGIRPVLGQEEVOFRLOALMECHDTRKEKRPALSLVSSQKMDPTFAT 240
DB 1506 HOLOIAKLSKSGIRPVLGQEEVOFRLOALMECHDTRKEKRPALSLVSSQKMDPTFAT 1565
QY 241 FMYELCCGKOTAFSSGOGERYTVFMDGKEESRNYTVNTEKGLMEVORRCCEGMKSCQ 300
DB 1566 FMYELCCGKOTAFSSGOGERYTVFMDGKEESRNYTVNTEKGLMEVORRCCEGMKSCQ 1625
QY 301 LOYORSIMTATEEDOKIYITYLKGMCPPLNTQOALDTPAVVTCFLAVVYIKKNSLYLAGL 360
DB 1626 LOYORSIMTATEEDOKIYITYLKGMCPPLNTQOALDTPAVVTCFLAVVYIKKNSLYLAGL 1646
QY 361 ADGLVAVFPVVRGTPKDCSYLCSHTANRSKFSIADEDARONPYPVKAMEVNSGSEVMY 420
DB 1647 ADGLVAVFPVVRGTPKDCSYLCSHTANRSKFSIADEDARONPYPVKAMEVNSGSEVMY 1706
QY 421 SNGPGLIYIDCASLEICRLEPYMAPSMYTSVYCSSEGRGEVYWCIDDKANSIYMYHST 480
DB 1707 SNGPGLIYIDCASLEICRLEPYMAPSMYTSVYCSSEGRGEVYWCIDDKANSIYMYHST 1766
QY 481 TYOLCARYFCGVSPPLDMFPVRPLDTEPPASHTANPKYPEGDSIADVSIMYSEELGTQ 540
DB 1767 TYOLCARYFCGVSPPLDMFPVRPLDTEPPASHTANPKYPEGDSIADVSIMYSEELGTQ 1826
QY 541 ILIHQESLTDYCSMSYSSSPPROAARSPSSLPSSPASSSVPESTCESDMLHTPGAA 600
DB 1827 ILIHQESLTDYCSMSYSSSPPROAARSPSSLPSSPASSSVPESTCESDMLHTPGAA 1886
QY 601 SDRSEHDLTPMDGETFSQHLOAVKILAVRDLIWPFRGGDYVYIGLEKDSGAQGRVIAY 660
DB 1887 SDRSEHDLTPMDGETFSQHLOAVKILAVRDLIWPFRGGDYVYIGLEKDSGAQGRVIAY 1946
QY 661 LKARELPHGVLDAAVAAVADTVCTEENETEMCLAVMKGAREDDIFYQSEELGRL 720
DB 1947 LKARELPHGVLDAAVAAVADTVCTEENETEMCLAVMKGAREDDIFYQSEELGRL 2006
QY 721 EACTRRK 727
DB 2007 EACTRRK 2013

RESULT 4

US-10-132-382-8
; Sequence 8, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132.382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2014
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-8

Query Match

Best Local Similarity 93.5%; Score 3574.5; DB 9; Length 2014;
Best Local Similarity 94.5%; Pred. No. 1.3e-274;
Matches 687; Conservative 0; Mismatches 1; Indels 39; Gaps 1;

QY 1 MLRLRATDAMKNSEFROEASMLHALQHPCTVALIGISHPCLFALELAPLSSINTVLS 60
DB 1327 MLRLRATDAMKNSEFROEASMLHALQHPCTVALIGISHPCLFALELAPLSSINTVLS 1386
QY 61 ENARDSSEFPLGMLTOKIAYQIASGLAYLHKNNIIFCDLKSNDIILWMSLDVKEHINKL 120

DB 1387 ENARDSSEFPLGMLTOKIAYQIASGLAYLHKNNIIFCDLKSNDIILWMSLDVKEHINKL 1446
QY 121 SDVIGISROSFHEGALGVEGPGYQAPETIRRIYDEKVDMSYGMVLYELLSCORPLGH 180
DB 1447 SDVIGISROSFHEGALGVEGPGYQAPETIRRIYDEKVDMSYGMVLYELLSCORPLGH 1506
QY 181 HOLOIAKLSKSGIRPVLGQEEVOFRLOALMECHDTRKEKRPALSLVSSQKMDPTFAT 240
DB 1507 HOLOIAKLSKSGIRPVLGQEEVOFRLOALMECHDTRKEKRPALSLVSSQKMDPTFAT 1566
QY 241 FMYELCCGKOTAFSSGOGERYTVFMDGKEESRNYTVNTEKGLMEVORRCCEGMKSCQ 300
DB 1567 FMYELCCGKOTAFSSGOGERYTVFMDGKEESRNYTVNTEKGLMEVORRCCEGMKSCQ 1626
QY 301 LOYORSIMTATEEDOKIYITYLKGMCPPLNTQOALDTPAVVTCFLAVVYIKKNSLYLAGL 360
DB 1627 LOYORSIMTATEEDOKIYITYLKGMCPPLNTQOALDTPAVVTCFLAVVYIKKNSLYLAGL 1647
QY 361 ADGLVAVFPVVRGTPKDCSYLCSHTANRSKFSIADEDARONPYPVKAMEVNSGSEVMY 420
DB 1648 ADGLVAVFPVVRGTPKDCSYLCSHTANRSKFSIADEDARONPYPVKAMEVNSGSEVMY 1707
QY 421 SNGPGLIYIDCASLEICRLEPYMAPSMYTSVYCSSEGRGEVYWCIDDKANSIYMYHST 480
DB 1708 SNGPGLIYIDCASLEICRLEPYMAPSMYTSVYCSSEGRGEVYWCIDDKANSIYMYHST 1767
QY 481 TYOLCARYFCGVSPPLDMFPVRPLDTEPPASHTANPKYPEGDSIADVSIMYSEELGTQ 540
DB 1768 TYOLCARYFCGVSPPLDMFPVRPLDTEPPASHTANPKYPEGDSIADVSIMYSEELGTQ 1827
QY 541 ILIHQESLTDYCSMSYSSSPPROAARSPSSLPSSPASSSVPESTCESDMLHTPGAA 600
DB 1828 ILIHQESLTDYCSMSYSSSPPROAARSPSSLPSSPASSSVPESTCESDMLHTPGAA 1887
QY 601 SDRSEHDLTPMDGETFSQHLOAVKILAVRDLIWPFRGGDYVYIGLEKDSGAQGRVIAY 660
DB 1888 SDRSEHDLTPMDGETFSQHLOAVKILAVRDLIWPFRGGDYVYIGLEKDSGAQGRVIAY 1947
QY 661 LKARELPHGVLDAAVAAVADTVCTEENETEMCLAVMKGAREDDIFYQSEELGRL 720
DB 1948 LKARELPHGVLDAAVAAVADTVCTEENETEMCLAVMKGAREDDIFYQSEELGRL 2007
QY 721 EACTRRK 727
DB 2008 EACTRRK 2014

RESULT 5

US-10-132-382-4
; Sequence 4, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132.382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2040
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-4

Query Match

Best Local Similarity 93.5%; Score 3574.5; DB 9; Length 2040;
Best Local Similarity 94.5%; Pred. No. 1.3e-274;
Matches 687; Conservative 0; Mismatches 1; Indels 39; Gaps 1;

QY 1 MLRLRATDAMKNSEFROEASMLHALQHPCTVALIGISHPCLFALELAPLSSINTVLS 60
DB 1353 MLRLRATDAMKNSEFROEASMLHALQHPCTVALIGISHPCLFALELAPLSSINTVLS 1412

QY 181 HOLOIAKLSKIGIRPYLGOPEEYQFRRLQALMECHDTEKREKRLALSVYSOKMDPTFAT 240
 DB 181 HOLOIAKLSKIGIRPYLGOPEEYQFRRLQALMECHDTEKREKRLALSVYSOKMDPTFAT 240
 QY 241 FMEELCCGKOTAFSSOGGEYTVFMDGKEESNRYTVNTEKLMVEQRMCCGMYSCQ 300
 DB 241 FMEELCCGKOTAFSSOGGEYTVFMDGKEESNRYTVNTEKLMVEQRMCCGMYSCQ 300
 QY 301 LOYORSIMATFEDOKIYITYLKMCPLNTPOALDTPAVVTCFLAPVYIKKNSYLVLAGL 360
 DB 301 LOYORSIMATFEDOKIYITYLKMCPLNTPOALDTPAVVTCFLAPVYIKKNSYLVLAGL 360
 QY 361 ADGLVAVFPVVRGTPKDSCSYLCSHTANRSKFSIADEDARQNPYPVKAMEVNSGSEVWY 420
 DB 361 ADGLVAVFPVVRGTPKDSCSYLCSHTANRSKFSIADEDARQNPYPVKAMEVNSGSEVWY 420
 QY 421 SNPGGLVTDCALEICRRLPEYMAPSMVTSVYCSSEGRGEYVWCLDCKANSLVYHST 480
 DB 421 SNPGGLVTDCALEICRRLPEYMAPSMVTSVYCSSEGRGEYVWCLDCKANSLVYHST 480
 QY 481 TYOLCARYCGVPSPLRDMFPVRPLDTEPPAASHTANPKYPEGDSIADYSIMYSEELGTQ 540
 DB 481 TYOLCARYCGVPSPLRDMFPVRPLDTEPPAASHTANPKYPEGDSIADYSIMYSEELGTQ 540
 QY 541 ILIHOSLTDYCSMSSYSSPPROAARSPSLPSSPASSSVFPSTDCEDSDMLHTPGAA 600
 DB 541 ILIHOSLTDYCSMSSYSSPPROAARSPSLPSSPASSSVFPSTDCEDSDMLHTPGAA 600
 QY 601 SDRSEHDLTPMGEETFSOHLQAVKILAVRDILVPRRGDVIYIGLEKSEAGRGVIAY 660
 DB 601 SDRSEHDLTPMGEETFSOHLQAVKILAVRDILVPRRGDVIYIGLEKSEAGRGVIAY 660
 QY 661 LKARELTIPRGVLYDAVAVAKDTVCTFENENTEMCLAVNRGMAAREFDIFYOSYEELGRL 720
 DB 661 LKARELTIPRGVLYDAVAVAKDTVCTFENENTEMCLAVNRGMAAREFDIFYOSYEELGRL 720
 QY 721 EACTRRR 728
 DB 721 EACTRRR 728

RESULT 2
 US-10-132-382-6
 ; Sequence 6, Application US/10132382
 ; Publication No. US20030045699A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 1987
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-132-382-6

Query Match 93.5%; Score 3574.5; DB 9; Length 1987;
 Best Local Similarity 94.5%; Pred. No. 1.3e-274;
 Matches 687; Conservative 0; Mismatches 1; Indels 39; Gaps 1;
 QY 1 MURHLRATDAMKNESEFROEASMLHALQHPCTVALIGISIHPLCFALDELAPSLSTLVLS 60
 DB 1300 MURHLRATDAMKNESEFROEASMLHALQHPCTVALIGISIHPLCFALDELAPSLSTLVLS 1359
 QY 61 ENARDDSSFIPLGHMLTOKIAYQIASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKL 120
 DB 1360 ENARDDSSFIPLGHMLTOKIAYQIASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKL 1419
 QY 121 SDYGISROSFHEGALGVEGTPTGQAPETIRPRIYDEKVMDFSTGMVLYELLGSGRALGH 180

DB 1420 SDYGISROSFHEGALGVEGTPTGQAPETIRPRIYDEKVMDFSTGMVLYELLGSGRALGH 1479
 QY 181 HOLOIAKLSKIGIRPYLGOPEEYQFRRLQALMECHDTEKREKRLALSVYSOKMDPTFAT 240
 DB 1480 HOLOIAKLSKIGIRPYLGOPEEYQFRRLQALMECHDTEKREKRLALSVYSOKMDPTFAT 1539
 QY 241 FMEELCCGKOTAFSSOGGEYTVFMDGKEESNRYTVNTEKLMVEQRMCCGMYSCQ 300
 DB 1540 FMEELCCGKOTAFSSOGGEYTVFMDGKEESNRYTVNTEKLMVEQRMCCGMYSCQ 1599
 QY 301 LOYORSIMATFEDOKIYITYLKMCPLNTPOALDTPAVVTCFLAPVYIKKNSYLVLAGL 360
 DB 1600 LOYORSIMATFEDOKIYITYLKMCPLNTPOALDTPAVVTCFLAPVYIKKNSYLVLAGL 1620
 QY 361 ADGLVAVFPVVRGTPKDSCSYLCSHTANRSKFSIADEDARQNPYPVKAMEVNSGSEVWY 420
 DB 1621 ADGLVAVFPVVRGTPKDSCSYLCSHTANRSKFSIADEDARQNPYPVKAMEVNSGSEVWY 1680
 QY 421 SNPGGLVTDCALEICRRLPEYMAPSMVTSVYCSSEGRGEYVWCLDCKANSLVYHST 480
 DB 1681 SNPGGLVTDCALEICRRLPEYMAPSMVTSVYCSSEGRGEYVWCLDCKANSLVYHST 1740
 QY 481 TYOLCARYCGVPSPLRDMFPVRPLDTEPPAASHTANPKYPEGDSIADYSIMYSEELGTQ 540
 DB 1741 TYOLCARYCGVPSPLRDMFPVRPLDTEPPAASHTANPKYPEGDSIADYSIMYSEELGTQ 1800
 QY 541 ILIHOSLTDYCSMSSYSSPPROAARSPSLPSSPASSSVFPSTDCEDSDMLHTPGAA 600
 DB 1801 ILIHOSLTDYCSMSSYSSPPROAARSPSLPSSPASSSVFPSTDCEDSDMLHTPGAA 1860
 QY 601 SDRSEHDLTPMGEETFSOHLQAVKILAVRDILVPRRGDVIYIGLEKSEAGRGVIAY 660
 DB 1861 SDRSEHDLTPMGEETFSOHLQAVKILAVRDILVPRRGDVIYIGLEKSEAGRGVIAY 1920
 QY 661 LKARELTIPRGVLYDAVAVAKDTVCTFENENTEMCLAVNRGMAAREFDIFYOSYEELGRL 720
 DB 1921 LKARELTIPRGVLYDAVAVAKDTVCTFENENTEMCLAVNRGMAAREFDIFYOSYEELGRL 1980
 QY 721 EACTRRR 727
 DB 1981 EACTRRR 1987

RESULT 3
 US-10-132-382-2
 ; Sequence 2, Application US/10132382
 ; Publication No. US20030045699A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEISS, BERTRAM
 ; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
 ; FILE REFERENCE: SCH-1811
 ; CURRENT APPLICATION NUMBER: US/10/132,382
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2013
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-132-382-2

Query Match 93.5%; Score 3574.5; DB 9; Length 2013;
 Best Local Similarity 94.5%; Pred. No. 1.3e-274;
 Matches 687; Conservative 0; Mismatches 1; Indels 39; Gaps 1;
 QY 1 MURHLRATDAMKNESEFROEASMLHALQHPCTVALIGISIHPLCFALDELAPSLSTLVLS 60
 DB 1326 MURHLRATDAMKNESEFROEASMLHALQHPCTVALIGISIHPLCFALDELAPSLSTLVLS 1385
 QY 61 ENARDDSSFIPLGHMLTOKIAYQIASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKL 120
 DB 1386 ENARDDSSFIPLGHMLTOKIAYQIASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKL 1445

Tue Apr 15 14:19:10 2003

us-09-836-392-21.rapb

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:25:34 ; Search time 38 seconds

(without alignments)
1171.238 Million cell updates/sec

Title: US-09-836-392-21

Sequence: 1 MNRHRAVDANKNFSEFROE.....IFQSYVELGRLEACTRRKR 728

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications-AA:*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3822	100.0	728	9	US-09-836-392-21
2	3574.5	93.5	1987	9	US-10-132-382-6
3	3574.5	93.5	2013	9	US-10-132-382-2
4	3574.5	93.5	2014	9	US-10-132-382-8
5	3574.5	93.5	2040	9	US-10-132-382-4
6	275.5	7.2	394	10	US-09-862-027-19
7	263	6.9	1036	12	US-10-014-882-2
8	260.5	6.8	847	9	US-10-143-133-2
9	243.5	6.4	505	9	US-09-977-260-6
10	243.5	6.4	505	9	US-09-977-261-6
11	243.5	6.4	505	10	US-09-977-269-6
12	243.5	6.4	505	10	US-09-982-610-20
13	243	6.4	455	10	US-09-757-982-5
14	242.5	6.3	850	10	US-09-904-389-2
15	238	6.2	937	9	US-09-974-298-129
16	237	6.2	425	10	US-09-828-313-29
17	236	6.2	263	10	US-09-840-704-5
18	235.5	6.2	966	10	US-09-771-161A-197
19	233	6.1	795	9	US-10-108-605-235

20	232.5	6.1	251	8	US-08-987-689A-30	Sequence 30, Appl
21	228.5	6.0	579	9	US-10-158-895-4	Sequence 4, Appl1
22	228.5	6.0	590	9	US-10-158-895-15	Sequence 15, Appl1
23	226	5.9	824	9	US-09-909-567B-53	Sequence 53, Appl
24	225	5.9	912	9	US-09-291-417-26	Sequence 26, Appl
25	225	5.9	968	9	US-09-291-417-107	Sequence 107, Appl
26	222	5.8	1036	10	US-09-771-161A-255	Sequence 255, App
27	222	5.8	1036	10	US-09-771-161A-256	Sequence 81, Appl
28	222	5.8	1734	10	US-09-862-027-81	Sequence 80, Appl
29	221.5	5.8	1265	10	US-09-862-027-80	Sequence 2, Appl1
30	220.5	5.8	752	10	US-09-835-081-2	Sequence 2, Appl1
31	220	5.8	1167	9	US-10-274-409-2	Sequence 10, Appl
32	218	5.7	265	10	US-09-797-039-10	Sequence 4, Appl1
33	218	5.7	278	9	US-09-842-582-4	Sequence 4, Appl1
34	218	5.7	278	9	US-09-934-406-4	Sequence 13, Appl
35	218	5.7	278	10	US-09-797-039-13	Sequence 17, Appl
36	218	5.7	278	10	US-09-922-138-27	Sequence 27, Appl
37	218	5.7	278	10	US-09-910-150-17	Sequence 31, Appl
38	218	5.7	278	10	US-09-910-150-31	Sequence 15, Appl
39	217.5	5.7	280	10	US-09-515-806-15	Sequence 16, Appl
40	217	5.7	278	10	US-09-815-915-16	Sequence 8, Appl1
41	216.5	5.7	835	10	US-09-947-199-8	Sequence 76, Appl
42	216.5	5.7	968	9	US-10-101-464A-76	Sequence 9, Appl1
43	215.5	5.6	459	10	US-09-740-627-9	Sequence 6, Appl1
44	215.5	5.6	459	10	US-10-172-088-6	
45	215	5.6	277	9	US-10-172-088-6	

ALIGNMENTS

RESULT 1
US-09-836-392-21
Sequence 21, Application US/09836392
Patent No. US20020173458A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptide
FILE REFERENCE: PRO2001
CURRENT APPLICATION NUMBER: US/09/836,392
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: PCT/US00/28066
PRIOR FILING DATE: 2000-10-11
PRIOR APPLICATION NUMBER: 60/159,542
PRIOR FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: 60/165,914
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: 60/189,027
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 21
LENGTH: 728
TYPE: PRT
ORGANISM: Homo sapiens
US-09-836-392-21
Query Match
Best Local Similarity 100.0%; Score 3822; DB 9; Length 728;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MNRHRAVDANKNFSEFROEASHALQPCIVMLGISIHPCLAPLAPLSLNTVLS 60
DB 1 MNRHRAVDANKNFSEFROEASHALQPCIVMLGISIHPCLAPLAPLSLNTVLS 60
OY 61 ENARDSFIPGHMLTORIAYQIASGLAYLTKRKNIFCDLSNDNLVWSLDVKRINKL 120
DB 61 ENARDSFIPGHMLTORIAYQIASGLAYLTKRKNIFCDLSNDNLVWSLDVKRINKL 120
OY 121 SDYGISRSFREGALVGEGTGGYQAPRIPIRYIDEKVDMSYGNVLYELLSCORPALGH 180
DB 121 SDYGISRSFREGALVGEGTGGYQAPRIPIRYIDEKVDMSYGNVLYELLSCORPALGH 180

DR	interPro: IPR001452; SH3.
DR	InterPro: IPR001245; Tyr_kinase.
DR	Pfam: PF00069; pkinase; 1.
DR	Pfam: PF00018; SH3; 1.
DR	PRINTS: PR00652; SH3DOMAIN.
DR	PRINTS: PD000001; Euk_kinase; 1.
DR	ProDom: PD000066; SH3; 1.
DR	SMART: SM00326; SH3; 1.
DR	SMART: SM00220; S_Terc; 1.
DR	SMART: SM00219; Tyrc; 1.
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR	PROSITE: PS50002; SH3; 1.
KW	kinase; SH3 domain.
FT	NON_TER
SO	SEQUENCE
	1066 AA; 118463 MW; EDD08BEEF7482723 CRC64;

Query Match	7.28;	Score 273.5;	DB 4;	Length 1066;
Best Local Similarity	31.38;	Pred. No. 6e-14;		
Matches	78;	Conservative 41;	Mismatches 93;	Indels 37;
				Gaps 8;

```

OY      3 RHLNATOMAKNFSEFROEASLTHALOPCIYALIGSIHP--LCFALIELPLSSINVLVS 60
Db      122 RHDEPDEISQIETENVRQEAFLKFAHLKPNRIATALGVCLKEPNCLYVMEFARGGELNVLVS 161
OY      61 ENARDSSEFPLGMLTOKI---AYQIASGLAYLHKK---NIIFCDIKSDNIIYWSL--- 110
Db      182 -----GKRIPPDILVMAVVOIARGNMYLHDEAIWPIIHRDLKSSNIILOKVEN 230
OY      111 -DYKEHINIKLSDGISROSFHEGALVEGPGVQADEIRPRIYDEKVMFSYGMVLYE 169
Db      231 GDLSSKRI-LKLTIDGELANHEHRTTKMSAAGTYANMAPEVIASFSKGSVDYGLIME 268
OY      170 LLSGQRPALGHNOLOIA----KLSKSGIRFVLGOPEYVORRLOALMMCEGMDTKPKRP 224
Db      290 LLTGEVPRGIDGLAVAGVAMNKLALPIPTCEP-----FAKIMEDCMNDPPHSRP 342
OY      225 LALSYYVSQM 233
Db      343 SFTNILDQL 351

```

Search completed: April 8, 2003, 10:25:30
Job time : 111 secs

RX NCBI_TaxID=4565;
 RN
 RP SEQUENCE FROM N.A.
 RA Niu J., Yu L., Chen P., Liu D., Ma Z.;
 RT "Molecular cloning and characterization of a serine/threonine protein
 RT kinase in wheat."; *Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.*
 RL EMBL: AF036609; AAK64576.1;
 DR EMBL: AF036609; AAK64576.1;
 DR InterPro: IPR000719; Ser_thr_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase.1.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN.1.
 DR ATP-binding; Kinase; Transferase.
 KW SEQUENCE 416 AA; 45958 MW; 54F986CED934353B CRC64;
 SQ
 Query Match 7.5%; Score 287; DB 10; Length 416;
 Best Local Similarity 35.0%; Pred. No. 1e-15;
 Matches 78; Conservative 38; Mismatches 91; Indels 16; Gaps 8;
 QY 16 EFRQASMLHALQPCVALIGISIHPL--CFALLETAPLSLMTVYSENARDSEFIPLGH 73
 Db 180 QFQGVYVMTALHNPVYKVGACRPIYWCITGYAKGGSVKNEL--NRQNRSVPL-- 235
 QY 74 MLOKIAVQIASGLAVLHKNNIFCDKSDNIWLSIDVKEHINIKLSIDYGISR-QSFHE 132
 Db 236 KLVKQALDVARGMAYVHGIFGRHDLKSDNLL-----ISGDKSIKIDAGVARIKTE 290
 QY 133 GALVGEVTPGYQAPETIRPRIVYDEKVDMSYGVAVYELLSSGGORPALGHOLQA-KKLK 191
 Db 291 GMPREGTIVKMAPEKQHPRIYKQKVDYISFGYVLMELLIGTLFPNNMTAVQAFAVAVK 350
 QY 192 GIRPVLGQPEEVOFRRLQALMECDTKPEKRPALSVSQM 234
 Db 351 GVRPAI--PHDC-LPALGEITRCMDANPVVRPFTDVVAMLE 390
 RESULT 14
 ID 022558 PRELIMINARY; PRT; 546 AA.
 AC 022558;
 DT 01-JAN-1998 (TReMBLrel.. 05, Created)
 DT 01-MAY-2000 (TReMBLrel.. 13, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel.. 20, Last annotation update)
 DE Hypothetical 61.5 kDa protein.
 GN AT2G17700.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Kounsley S.D., Shua T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Rongning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Varakan S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."; *Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.*
 RL Nature 402:761-768 (1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Kounsley S.D., Shua T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Rongning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Varakan S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."; *Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.*
 RL Nature 402:761-768 (1999).
 RN (3)
 RP SEQUENCE FROM N.A.

RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carlini P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Lam B., Lin J., Meyers H.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full length cDNA of gene At2g17700 (GI:15227883).";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF024504; AAB80785.2;
 DR EMBL: AY070086; AAL49781.1;
 DR HSP: P08631; IAD5.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR Pfam: PF01842; ACT.1.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk_pkinase.1.
 DR SMART: SM00221; STYKc.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;
 KW Transferase.
 SQ SEQUENCE 546 AA; 61509 MW; E74F30F60D91F426 CRC64;
 Query Match 7.3%; Score 278.5; DB 10; Length 546;
 Best Local Similarity 31.8%; Pred. No. 8.3e-15;
 Matches 71; Conservative 45; Mismatches 90; Indels 17; Gaps 8;
 QY 16 EFRQASMLHALQPCVALIGISIH--PLCFALLETAPLSLMTVYSENARDSEFIPLGH 73
 Db 327 EFSQEVFMKRVKKNVOPLGACTNSPTLCITYEFNARSTYDFEL--KQCAEFLQT 383
 QY 74 MLOKIAVQIASGLAVLHKNNIFCDKSDNIWLSIDVKEHINIKLSIDYGISRQSFHEG 133
 Db 384 LL--KVALDVAKMSYLNHNNIHRDLKTANLL-----MDEHGLVYKADGVARIQESG 436
 QY 134 ALVGEVTPGYQAPETIRPRIVYDEKVDMSYGVAVYELLSSGGORPALGHOLQA-KKLK 191
 Db 437 VMTAEIGTYKMAPEVYEHKPNYKADVFESYAVYELMELLTGDIPIYAFELPLQAAVGVAK 496
 QY 192 GIRPVLGQPEEVOFRRLQALMECDTKPEKRPALSVSQM 234
 Db 497 GLRPKI--PKTH-PKVGLIERCHMODPQRPLFEETIMLQ 536
 RESULT 15
 ID 09H2N5 PRELIMINARY; PRT; 1066 AA.
 AC 09H2N5;
 DT 01-MAR-2001 (TReMBLrel.. 16, Created)
 DT 01-JUN-2001 (TReMBLrel.. 16, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel.. 21, Last annotation update)
 DE Mixed lineage kinase MK1 (Fragment).
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC McNea J.J., Dower S.K., Guesdon F.;
 RT "cDNA sequence and gene organisation of mixed lineage kinase 1."; *Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.*
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL: AF251442; AAG44591.1;
 DR HSP: P29355; ISEM.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_kinase.

Db 387 E-----SIACSKFTLPQESFHYMCWKRKEIAGVLHDLREKERE--LRNKEBQL 435
 QY 286 EVQ 288
 Db 436 RVQ 438

RESULT 11

Q95VF6 PRELIMINARY; PRT; 1148 AA.

AC Q95VF6;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Mixed lineage kinase
 GN MLK2 OR DMLK OR CG2272.
 OS Drosophila melanogaster (Fruit Fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sathyanarayana P., Tzivion G., Barthwal M.K., Rana A.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF416233; AAI08011.1;
 DR FlyBase: FBgn0030018; MLK2.
 DR Interpro: IPR000719; Euk_Pkinase.
 DR Interpro: IPR002290; Ser_thr_Pkinase.
 DR Interpro: IPR001452; SH3.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Prodom: PD000001; Euk_Pkinase; 1.
 DR Prodom: PD000066; SH3; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
 DR PROSITE: PS50002; SH3; 1.
 KW ATP-binding; Kinase; Transferase.
 SQ SEQUENCE 1148 AA; 127511 MW; 66A7D7A6E75E116C CRC64;

Query Match 7.5%; Score 288.5; DB 5; Length 1148;
 Best Local Similarity 29.0%; Pred. No. 3.6e-15;
 Matches 88; Conservative 58; Mismatches 120; Indels 37; Gaps 12;

QY 2 LRHLRATPAMKNFSE-FQGEASMLHALQPCVALIGISIH-PLCFALFLAPLSLNTVL 59
 Db 157 IAHOTGEDDMQRMNDVNLQEAFLKFWALKHENIAALRGVCLNTKICLVMEYARGSLNRL 216
 QY 60 SENARDSFIPLGMLTKIAYQIASGLAYLHK---NIIFCDKSDNIIWVSLDKVKEHI 116
 Db 217 AGK-----IPPDVLVNM--AIQIARGMNYLHNEAPMSIIHRDLKSSNVLLYEIEGNHL 268
 QY 117 ---NIKISDYGISROSFEHAGLVGEGTPGYQAPETIRPIVYDEKVDMSYGMVLEYELLSG 173
 Db 269 QOKTKIKIDFGIARMTYQMSAAGTYAMPPEVISTYSKFSQVSYGLVLMELLITG 328
 QY 174 QRPALGHQLOIAKKSIGIRPVLGQPEVOFRRLQALMECWDTPKPKRPLASVVSOM 233
 Db 329 ETPYKGFPLSVAYGVAVNTL-TLPPIKTCF-ETWGALMKSCWQTDPRKRPGEKILKOL 386
 QY 234 KDPFATPMYELCKCK-----QTAFSSQ---GOEYTVFMDGKESNRYTVVNTKEKLM 285
 Db 387 E-----SIACSKFTLPQESFHYMCWKRKEIAGVLHDLREKERE--LRNKEBQL 435
 QY 286 EVQ 288
 Db 436 RVQ 438

RESULT 12
 Q95UN8 PRELIMINARY; PRT; 1161 AA.
 AC Q95UN8;

DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Mixed lineage protein kinase.
 GN MLK2 OR SLPR OR CG2272.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-HEAD;
 RA Stronach B.E., Perrimon N.;
 RT Activation of the JNK pathway during dorsal closure in Drosophila
 RT requires the mixed lineage kinase, slapper.
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY045717; AAR98795.1;
 DR FlyBase: FBgn0030018; MLK2.
 DR Interpro: IPR000719; Euk_Pkinase.
 DR Interpro: IPR002290; Ser_thr_Pkinase.
 DR Interpro: IPR001452; SH3.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Prodom: PD000001; Euk_Pkinase; 1.
 DR Prodom: PD000066; SH3; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
 DR PROSITE: PS50002; SH3; 1.
 KW ATP-binding; Kinase; Transferase.
 SQ SEQUENCE 1161 AA; 128960 MW; DD6C7ABAC08EDA24 CRC64;

Query Match 7.5%; Score 288.5; DB 5; Length 1161;
 Best Local Similarity 29.0%; Pred. No. 3.7e-15;
 Matches 88; Conservative 58; Mismatches 120; Indels 37; Gaps 12;

QY 2 LRHLRATPAMKNFSE-FQGEASMLHALQPCVALIGISIH-PLCFALFLAPLSLNTVL 59
 Db 170 IAHOTGEDDMQRMNDVNLQEAFLKFWALKHENIAALRGVCLNTKICLVMEYARGSLNRL 229
 QY 60 SENARDSFIPLGMLTKIAYQIASGLAYLHK---NIIFCDKSDNIIWVSLDKVKEHI 116
 Db 230 AGK-----IPPDVLVNM--AIQIARGMNYLHNEAPMSIIHRDLKSSNVLLYEIEGNHL 281
 QY 117 ---NIKISDYGISROSFEHAGLVGEGTPGYQAPETIRPIVYDEKVDMSYGMVLEYELLSG 173
 Db 282 QOKTKIKIDFGIARMTYQMSAAGTYAMPPEVISTYSKFSQVSYGLVLMELLITG 341
 QY 174 QRPALGHQLOIAKKSIGIRPVLGQPEVOFRRLQALMECWDTPKPKRPLASVVSOM 233
 Db 342 ETPYKGFPLSVAYGVAVNTL-TLPPIKTCF-ETWGALMKSCWQTDPRKRPGEKILKOL 399
 QY 234 KDPFATPMYELCKCK-----QTAFSSQ---GOEYTVFMDGKESNRYTVVNTKEKLM 285
 Db 400 E-----SIACSKFTLPQESFHYMCWKRKEIAGVLHDLREKERE--LRNKEBQL 448
 QY 286 EVQ 288
 Db 449 RVQ 451

RESULT 13
 Q94C42 PRELIMINARY; PRT; 416 AA.
 AC Q94C42;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Serine/threonine protein kinase.
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.

```

DB 262 DDICNKTITDFGLAREHRTTBMASACTYAMMAPEVIRSSLSFGSDIMSGVILML 321
OY 171 LSGORPALGHOLQTA-----KLSGICIPVJGQPEVOPFRRLQALMECWDTPKRP 225
DB 322 LGEVFPYRIGDILAAYGVAVKRLPLIPSTCEP-----FAKLMKCEWEDPHIR-- 372
OY 226 ALSVMSOMDPFATFMEYELCCGKOTAFPSGOGEXTVVFMDKESRNYVTNT-ERKL 284
DB 373 -----PSFALILQOULTAIEEAVLTNMQESFHSQEDWKLEIQFSELRKEKEL 422
OY 285 M-----EVORMCPGKMKVSCOLO-----VORSIMTATEDOKIYITLKMCPILNPOOL 334
DB 423 RSREELSAALQOKSQELLRRREQQLAREIDVLERLANLITOLS-----QEA- 473
OY 335 DTPAVVTCCLAVPVIKKNSYLVLG-----LADGLVAVP-----VVRGTP-----KDCS 379
DB 474 -----PHYKRRGRFRGRGLRLKDGHRISLSPDFQHKITVQASPLDKRRSSD 521
OY 380 SYLCSHTAN-----RSKFSIADSD-----ARQNPYVKAMEVNSGSEVWYNSG 423
DB 522 SCLSPSPGSPMLPLRLRALQTLSDENKTRGNMVFROEDFE-----DYKRSKRRKGGCTWG 577
OY 424 PGLIVIDCASLEICRLEFYMAPSMVTSYVSCSESGEGEEVWCILD-----KANSLVYHST 480
DB 578 P-----SSVQTKEREPEGRRVRVPLDGNPSMSLILKSK 612
OY 481 TYQOLARFCGVPSPLRDMFPV-----RPIDTEPPAASHTANRV----- 520
DB 613 TTPPL-ASLFEVDPGSCCEQKLVPEGLERHKPKOTKFGCAHVLGLPKDQSDSEASES 671
OY 521 -----PEGSIDVDS-----IMYSEELGTQI-LIH----- 544
DB 672 REEGSPKSPVNVGAPMIRKRTESALCEGMLASMLADVRKLHGQAQAPAKSPKXE 731
OY 545 ---QESLTDYCSMSYSSSPRO--AARSPS-----IPSSPASSSVYFTDC----- 588
DB 733 KKEGALPASHRCQSSPSLLRQPSAGRAPSGSSTLLPLSPASVSHSKSLMKCLLQAGK 791
OY 589 EDSMDLHT-----PGAASDRSHDITP 610
DB 792 EESSLGNADLCGPTTLTPPDGSAPESSCCLIP 825

RESULT 10
O9W313 PRELIMINARY: PRT: 1020 AA.
AC 09W313:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CG2272 protein.
GN MK2 OR CG2272.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_Taxid=7227;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yeaman M.D., Zhang Q., Chen L.X.,
RA Brindon R.C., Rogers J.-H.C., Blazek R.G., Champs M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.N., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burdits K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

```

```

RA Cherry J.M., Cavley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Doudon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Duthin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaser K.,
RA Foster A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mel M.-H., Ibegwan C.,
RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laekso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei P., McIntosh T.C., McLeod M.P., Metherson D.,
RA Mikhlov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon R., Nusser D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sledge-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL: AE003443; AAF6344.1.
DR HSSP: P11362; ITCR.
DR FlyBase: FBgn0030018; MK2.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_thr_Pkinase.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR004040; Ser_Pkinase.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_Pkinase; 1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00326; SH3; 1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00002; SH3; 1.
KW ATP-binding; SH3 domain; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1020 AA; 11370 MD; D48BD9C6639A08F CRC64;

Query Match 7.5%; Score 288.5; DB 5; Length 1020;
Best Local Similarity 29.0%; Pred. No. 3e-15;
Matches 88; Conservative 58; Mismatches 120; Indels 37; Gaps 12;

OY 2 LRLHATAMNFSE-FQESAMTALHPCLVALIGSIH-PLCFALAPLSLNTVL 59
DB 157 IAHOTGEDMDORMDNVIOEAKLFALRHENIAALRGVCLNKLCTLVNEYARGSLNRL 216
OY 60 SENARDSFPIGHLTKIVQINSGLAYLHK--NIFQDLKSDMLVMSLDVRHI 116
DB 217 AGK-----IPDVIVNV--AIQIRGMNYLHNEPMSIHHDLSSNVLLITAEIGHNL 268
OY 117 --NIKLSDYISROSFEHALGVEGTPGYAPETRIYDEKVDMSYGVLYELSG 173
DB 269 QOKTLKINDFLGARMYWTQBSAAGTVAWMPFVIVSIVTSKSFSDVSYGLMLLITG 328
OY 174 QPRLAGHOLQIAKLKSGIPVJGQPEVOPFRRLQALMECWDTPKRPALASVNSG 233
DB 329 ELPKYGFPLSVAAVAVNTL-TLPITPTCP-ETWGAIAKRSQWQTDPPKRGFXELKQL 386
OY 224 KDPFATWYELCCGK--QIAFFSSO--CGEYTVVFNQDEESNNTYVNTKGLM 285

```

DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; kinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk_kinase.1.
 DR SMART: SM00221; STYKC.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS50018; PROTEIN_KINASE_ST.1.
 DR APP-binding_kinase: Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 412 AA; 46083 MW; 644F35A90210D48 CRC64;
 SO

Query Match 8.2%; Score 315; DB 10; Length 412;
 Best Local Similarity 35.5%; Pred. No. 4,4e-18;
 Matches 82; Conservative 44; Mismatches 73; Indels 32; Gaps 9;

QY 16 EFRQASMLALHPICVALIGISIHPL--CFALELAPLSLNTLVSENRDSSFFPLGH 73
 DB 176 QFOQEVSMALAKHPNIYRFAGACIRPMWCIVTEYAKGSSVROFLTR--RQNRVPLKL 233
 QY 74 MLTKRIAYQASGLAYLHKNNIIFCDKSDNIIWMSLDVKEHINKIKSDYGISRQSFH-E 132
 DB 234 AVMG--ALDVARGMAVYHGRNFIRHDKSDMLTI-SADK---SRIKIDFVARIEVQTE 286
 QY 133 GALVEGTPGYQAEIPRIYDEKVMFSYGMVLEYELSGORALGHQOIA-KRLSK 191
 DB 287 GMPETGTYRMAPEMIOHRVYKVDYISGVIYLMELITLLEPQNTAVQAAPAVYNR 346
 QY 192 GIR-----PVLGQPEEVOFRRLQALMECWDTRPEKRPALSVSQM 234
 DB 347 GVRPTVPADCLPVLGE-----ITRCWDADPEVRPCFEIVALLE 386

RESULT 8
 Q92031 PRELIMINARY; PRT; 411 AA.
 AC Q92031; Q94A10;
 DT 01-MAY-1999 (TRENBLREL. 10, Created)
 DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Hypothetical 46.0 kDa protein.
 CN ATG24360.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanaken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nigam W.C.,
 RA Fraser C.M., Venter J.C., Buell C.R., Town C.D., Nigam W.C.,
 RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Gibson H.A.,
 RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Yu G., Yu S., Bowser L., Carinci P., Chen H.,
 RA Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamita A.,
 RA Kallin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southlick A., Tracy S.E.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.; Tracy S.E.,
 RT "Full length cDNA of gene T28124.9/At2g24360 (GI:337195)."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006403; AAD18109.2;
 DR EMBL: AY046026; AAK76700.1;
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002390; Ser_thr_kinase.
 DR Pfam: PF00069; kinase.1.
 DR PRODOM: PD000001; Euk_kinase.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS50018; PROTEIN_KINASE_ST; UNKNOWN.1.

KW Hypothetical protein; ATP-binding; Transferase.
 SO SEQUENCE 411 AA; 46001 MW; 3B7001CB841BAA CRC64;
 SQ

Query Match 8.2%; Score 314; DB 10; Length 411;
 Best Local Similarity 36.8%; Pred. No. 5.3e-18;
 Matches 82; Conservative 42; Mismatches 83; Indels 16; Gaps 9;

QY 16 EFRQASMLALHPICVALIGISIHPL--CFALELAPLSLNTLVSENRDSSFFPLGH 73
 DB 175 QFOQEVSMALAKHPNIYRFAGACIRPMWCIVTEYAKGSSVROFLTR--RQNRVPLKL 230
 QY 74 MLTKRIAYQASGLAYLHKNNIIFCDKSDNIIWMSLDVKEHINKIKSDYGISRQSFH-E 132
 DB 231 KLAVQALDVARGMAVYHGRNFIRHDKSDMLTI-SADK---SRIKIDFVARIEVQTE 285
 QY 133 GALVEGTPGYQAEIPRIYDEKVMFSYGMVLEYELSGORALGHQOIA-KRLSK 191
 DB 286 GMPETGTYRMAPEMIOHRVYKVDYISGVIYLMELITLLEPQNTAVQAAPAVYNR 345
 QY 192 GIRPVLGQPEEVOFRRLQALMECWDTRPEKRPALSVSQM 234
 DB 346 GVRPTV--PNDCLPVLSDINTRCWDANPEVRPCFEIVALLE 385

RESULT 9
 Q8VDG6 PRELIMINARY; PRT; 1001 AA.
 AC Q8VDG6;
 DT 01-MAR-2002 (TRENBLREL. 20, Created)
 DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Similar to mitogen-activated protein kinase kinase 9.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strauberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC021891; AAH21891.1;
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002390; Ser_thr_kinase.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; kinase.1.
 DR Pfam: PF00018; SH3.1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRODOM: PD000001; Euk_kinase.1.
 DR SMART: SM00326; SH3.1.
 DR SMART: SM00220; S_TKC.1.
 DR SMART: SM00219; TYKIC.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS50018; PROTEIN_KINASE_ST; UNKNOWN.1.
 DR PROSITE: PS50002; SH3.1.
 KW kinase
 SO SEQUENCE 1001 AA; 109983 MW; E10042C86B9953C CRC64;
 SQ

Query Match 7.6%; Score 292; DB 11; Length 1001;
 Best Local Similarity 22.8%; Pred. No. 1.5e-15;
 Matches 172; Conservative 99; Mismatches 257; Indels 226; Gaps 32;

QY 9 DAMKFSERQASMLALHPICVALIGISIHPL--CFALELAPLSLNTLVSENRDSSFFPLGH 66
 DB 146 DAAAAESVRRERARLFAMLRHPIITOLRGVCLRPOLCLIVLEFAGGALNALAAADP 205
 QY 67 SFIPIGHMLTKI-----AYQIASGLAYLHK--NIIFCDKSDNIIWMSLDVKEH 115
 DB 206 R--APGPRARRIPQVLYNMAVQIARGLILHEAAVPIILHRDKSSNILL--LEKIEH 261
 QY 116 INI-----KLSDYGISRQSFHFGALGVGTPGYQAEIPRIYDEKVMFSYGMVLEYEL 170

RA Amanatides P.G., Scheer S.E., Li P.W., Hoskins R.A., Galle R.E.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Handerson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Barker E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Abayaratne A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Bailey R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Bereman B.P., Bhanderi D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jellali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Palenart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Sprengel A.C., Stapleton M., Strong R., Sun E.,
 RA Svistskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 "The genome sequence of *Drosophila melanogaster*,"
 Science 287:2185-2195(2000).
 RL -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC EMBL; AB003731; AAF55793.1; -
 DR P1BASE; FBgn0038816; CG5483.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_in.
 DR InterPro: IPR001806; Ras_tnsfmg.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00023; ank; 7.
 DR Pfam: PF00060; LRR; 9.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART; SM00248; ANK; 3.
 DR SMART; SM00370; LRR; 5.
 DR SMART; SM00369; LRR_TYP; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 1.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR ANK repeat: ATP-binding; GTP-binding; Repeat;
 KW Serine/threonine-protein kinase; Transferrin;
 SQ SEQUENCE 2308 AA; 256468 MW; CE36F007E79D5D0C CRC64;

Query Match 9.5%; Score 364.5; DB 5; Length 2308;
 Best Local Similarity 23.7%; Pred. No. 3,9e-21;
 Matches 159; Conservative 99; Mismatches 243; Indels 169; Gaps 27;

QY 2 LNLRLATDAKMFSEFRQASMLHALQHCIALGICSHIPCFALFALSLNLTVE 61
 DB 1684 LGH-----SCRAYCARQAGLAVLTLTKHKNITPVIVGICIKPLAVIELAPLGGLDLHLH 1738
 QY 62 NARDSSTPLGMLTOKIAYQIASGLAYLHKNNITFCDKSNITLWML-----DYKE 114

DB 1739 YRRSGAH--MGPHRTQTLVLAARALEYLRHRRRIITRDLSBNVLTWELPOPHTEDESPN 1796
 QY 115 HINIKLSDYIRSGFHEGALVEGTPGYQAEIRPRIYDEKVDMEFSGVLTLLSQ 174
 DB 1797 LVHIKADIGIRQAPSAKAFGGTEGMAPEI---IKYN----- 1834
 QY 175 RPALGHHQIAKLSKIGIRPVLAGPEEYQFR--LQALMEGCDKREKRPALSYVSCM 233
 DB 1835 -----GHE--SIKECILTEGSRPALTO--RETQFPCCIDLAVLWLMHQPPRRPASAQVSTL 1887
 QY 234 KRPFRATEH-----YELCGCK-QTAFSSGGEITYVFPDGRKESNNYVNTKES- 283
 DB 1888 SAPECIHLIDVYAMPSEKIVGVQSLVGMDDERCGLMLPSRGSIDILDCSPSGS 1947
 QY 284 LMEVQRMCC-----PGKVSQLOVQRSLMTATEDQRY 317
 DB 1948 LIQCSISCSPPQYAPPTPENGANSRARSQRPLPKMMLCCCLVGEALIMGDVSGNLIH 2007
 QY 318 IYTLGMCPLNTPQOALDTPAVVTCFLAVPKNSYLVLAGLADGLVAVFPVNGTPPD 377
 DB 2008 AVSTSTYAHIFS--YMLD-PNLSAVISLVYMEKIA-RVAVGTNG--RVFLVDATQMS 2061
 QY 378 SCST-----LCSHTANRSKFSIADSD-----ANQNPYP----- 405
 DB 2062 NCAPFEGSEFVLEICSGFVLAACSVVDGIYELMCGEYIAGKINFPPLNENGVSQHQLC 2121
 QY 406 -----VAMENVNGSEVWVNSNGGLV-----IDCASLEIC-RILE 441
 DB 2122 HSEEPNLIEDVYVARRCSHSHFSCLYPGCMYQMDYISKRIENKLDCKSLPSESIQ 2181
 QY 442 PYMAPSMYTVSVCSSEGRGEVYWCIDKANSLVWYHSTYOLCARFYCGVSPPLRMDP 501
 DB 2182 STAIDENHVLIC-----QISLAHNSLEY-IGTTWGLIYAEHLTLRP 2225
 QY 502 V--RPDTEPPAASHTANPKVPEGDSIADVSIMSEELGTOILHQSLETPYCSMSYS 558
 DB 2226 ISVFPREYNEIKSITLSKNDVP---LNTIGRRY-RSLISRYVDSAEEST---KSSAV 2277
 QY 558 SSPPQARS 568
 DB 2278 STPHGAARS 2287

RESULT 7
 Q9M085 PRELIMINARY; PRT; 412 AA.
 AC Q9M085;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Protein kinase-like protein.
 GN Atg631170.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI-TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lennard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.,
 RA Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL161578; CAB99835.1; -
 DR HSP; P08631; IAD5
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR004040; STY_pkinase.

RESULT 4
Q9BS11 PRELIMINARY; PRT; 221 AA.
AC Q9BS11; ID Q9BS11; DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Similar to hypothetical protein FLN23119.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC Tissue-Cervix;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC005408; ANH05408.1; -
KW Hypothetical protein.
SQ SEQUENCE 221 AA; 24141 MW; ABC2413CC4C595E CRC64;
Query Match 27.0%; Score 1031; DB 4; Length 221;
Best Local Similarity 85.3%; Pred. No. 4.6e-79;
Matches 203; Conservative 6; Mismatches 9; Indels 20; Gaps 3;
QY 476 MYHSTYQLCARYFCGVPSPFLDMFVPLDTEPPASHTANPKVPEGDSIADVSIMYSE 535
DB 1 MYHSTYQLCARYFCGVPSPFLDMFVPLDTEPPASHTANPKVPEGDSIADVSIMYSE 60
QY 536 ELTQTLIHESLTDYCSMSYSSPPROAARSPSSLPSSPASSSVPESTDCDSMLH 595
DB 61 ELTQTLIHESLTDYCSMSYSSPPROAARSPSSLPSSPASSSVPESTDCDSMLH 120
QY 596 TPGAASRSEHDLTPMDGFEFFSOHLQAKILAVNDLWVRGGDVYIGLEKSEKQRG 655
DB 121 MPGAASRSEHDLTPMDGFEFFSOHLQAKILAVNDLWVRGGDVYIGLEKSEKQRG 180
QY 656 RVIAVLKARELTPHGVLDVAAYVAKDTVCTFENENTWCLAVWRGNAREFDITFYOS 713
DB 181 RVIAVLKARELTPHGVLDVAAYVAKDTVCTFENENTWCLAVWRGNAREFDITFYOS 218
RESULT 5
Q9TZM4 PRELIMINARY; PRT; 1175 AA.
AC Q9TZM4; ID Q9TZM4; DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical 130.7 kDa protein.
GN T27C10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Znu H.J., Graves T., Hawkins M.;
RT "The sequence of C. elegans cosmid T27C10.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF098504; AAK73897.1; -
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Hypothetical protein; Transferase.
SQ SEQUENCE 1175 AA; 130667 MW; F4BFBF44DDC1390C CRC64;
Query Match 10.5%; Score 402; DB 5; Length 1175;
Best Local Similarity 27.7%; Pred. No. 9.5e-25;
Matches 123; Conservative 86; Mismatches 175; Indels 60; Gaps 19;
QY 10 AMKNESEFROEASMLHALHPCIVALLGISHPCLALELAPLSLNTVUSENARSSFT 69
DB 542 ACRACTSRQELSLSMKHPNVIGVCTFPLSLVELAPLAGALNQLGSHRKAGTKL 601
QY 70 PLGHMLTKIAVOIASGLAVLHKKNITFCDLKSDNLIWSDV---KEHINIKLSGYGI 125
DB 602 SLG--VIKESAVOVARALEYLSAHNITRDKSENVLGMNRPAPFSPQDVLTKLDYGI 659
QY 126 SROSFHE-GALGVEGTPGYAPEI---RPRIYDEKVMDSYGMVLYELLSGORPALGH 181
DB 660 SRVLPSGAKGFGGTGFAPAEIVRNGEETQKVDCSFGMLYELLTKFPF--ES 717
QY 182 QLOIAKLSGIRPVLSQPEEVOFRRIQALMECMTDTPKRPALASVYQMKDPPTATF 241
DB 718 EEHVKERMLDGAAPVLLPHELLPTPMLDLVHCWSAHPSRSSSQLVQFCAPAF-T 776
QY 242 MYELC-CKG---QTAFPS-----SOGGYTVFMGKRESNRYTVN--TEKGMEYOR 289
DB 777 LLDVCELGEMLPFQLMNAVITDEIDPDPEAQLM---LSGEMVYMGCTQVGFQDKS 833
QY 290 MCEP--GMKYSOLOVORSMTATEDDKIIYTLTKGMCPLNTPQALDTPAVVTCFLAVP 347
DB 834 IELPHRKYYVS---KVADSVWSCDECGQVYV---GSLHETGH--LQLPSLNTLICAP 885
QY 348 VIKNSTLYIAGLADGVAVFPPVVRGTPKDCSTLCSHTANRKSFTIADDAQNPPVVK 407
DB 886 ELISNDVLLI--ISDKQIVLLKL---SESNVSHL-----GTIDSPYEIR 925
QY 408 AMEVNCSG--EWMYNGPGLVI 429
DB 926 TATFLNGSTROIWAGHSEGRISI 949
RESULT 6
Q9VDJ9 PRELIMINARY; PRT; 2308 AA.
AC Q9VDJ9; ID Q9VDJ9; DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE CG5483 protein.
GN CG5483.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

Db 642 MRRHRTADAMNFESEFQREASMLHLPICVIALGISIHPLCEFLAPLSLNTVLS 701
Oy 61 ENARSSFIPLGMLTQKLAIOIASGLAYLHKNNIFCDLSDNLIWMSLDVREHNITL 120
Db 702 ENARSSFIPLGMLTQKLAIOIASGLAYLHKNNIFCDLSDNLIWMSLDVREHNITL 761
Oy 121 SDYGISROSFHGALGVECTPGYQAPETPRIVYDEKVMFSYGVLYELLSGORPALGH 180
Db 762 SDYGISROSFHGALGVECTPGYQAPETPRIVYDEKVMFSYGVLYELLSGORPALGH 821
Oy 181 HOLQAKKLSKGIKIPVGLGQPEEQFRRLQALMECHDTPERKPLALSVSQMKDPTFAT 240
Db 822 HOLQAKKLSKGIKIPVGLGQPEEQFRRLQALMECHDTPERKPLALSVSQMKDPTFAT 881
Oy 241 FMYELCCGKOTAFSSGOGELYTVWMDGKEESRNTVYVTEKGLMEVQRMCKPGMKVSCQ 300
Db 882 FMYELCCGKOTAFSSGOGELYTVWMDGKEESRNTVYVTEKGLMEVQRMCKPGMKVSCQ 941
Oy 301 LOVQSLMTATEDOKIYITLTKGKPLNTPOALDTPAVVTCFLAVPVIKKNSYLVLAGL 360
Db 942 LOVQSLMTATEDOKIYITLTKGKPLNTPOALDTPAVVTCFLAVPVIKKNSYLVLAGL 1001
Oy 361 ADGLVAVFPVVGTPKDCSYLCSHTANRSKFSIADEDARQNPYPKAMEVNVSSGEVWY 420
Db 1002 ADGLVAVFPVVGTPKDCSYLCSHTANRSKFSIADEDARQNPYPKAMEVNVSSGEVWY 1061
Oy 421 SNGPGLIVIDCASLEICRLREPYMAPSMTSVYCSSESGEVEVWCLDKANSIYVYHST 480
Db 1062 SNGPGLIVIDCASLEICRLREPYMAPSMTSVYCSSESGEVEVWCLDKANSIYVYHST 1121
Oy 481 TYQOLARFCGVPSPRLDMFPVRLDTEPPAASHTANPKYPEGDSIADVSIMYSEELGTQ 540
Db 1122 TYQOLARFCGVPSPRLDMFPVRLDTEPPAASHTANPKYPEGDSIADVSIMYSEELGTQ 1181
Oy 541 IILHOESLTDYCSMSYSSSPROAARSPSSLPSSPASSSVPFSDCCSDMLHTPGAA 600
Db 1182 IILHOESLTDYCSMSYSSSPROAARSPSSLPSSPASSSVPFSDCCSDMLHTPGAA 1241
Oy 601 SDRSEHDLTMDGETFSQHLQAVKILAVRDLIMVPRGGDIVIGLEKDESAQRGVIVAY 660
Db 1242 SDRSEHDLTMDGETFSQHLQAVKILAVRDLIMVPRGGDIVIGLEKDESAQRGVIVAY 1301
Oy 661 LKARELTPHGVLDVAAYVAKDTVCTFENENTEMCLAVMRGAGAREDFIYQSYEELGL 720
Db 1302 LKARELTPHGVLDVAAYVAKDTVCTFENENTEMCLAVMRGAGAREDFIYQSYEELGL 1361
Oy 721 EACTRRKR 728
Db 1362 EACTRRKR 1369

```

RESULT 2

```

OyH5S3 PRELIMINARY: PRT: 253 AA.
AC 09H5S3:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CDNA: FLJ23119 fls, clone LNG07978.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Ox NCBI_TaxID=9606;
RN [1].
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T., Matsumura K.,
RA Nakaajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Ohbayashi M., Nishii T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isegaki T., Sugano S.;
RT "NEDO human cDNA sequencing project."

```

Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK026772; BAB15547.1;
 SQ SEQUENCE 253 AA; 27913 MW; FAFD58E9E29FE741 CRC64;

Query Match 34.7%; Score 1325; DB 4; Length 253;
 Best Local Similarity 99.6%; Pred. No. 7.7e-104;
 Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy 476 MYHSTTYOLCARFCGVPSPRLDMFPVRLDTEPPAASHTANPKYPEGDSIADVSIMYSE 535
Db 1 MYHSTTYOLCARFCGVPSPRLDMFPVRLDTEPPAASHTANPKYPEGDSIADVSIMYSE 60
Oy 536 ELGTQILIHQESLTDYCSMSYSSSPROAARSPSSLPSSPASSSVPFSTDCEDSDMLH 595
Db 61 ELGTQILIHQESLTDYCSMSYSSSPROAARSPSSLPSSPASSSVPFSTDCEDSDMLH 120
Oy 596 TPGAASDRSEHDLTMDGETFSQHLQAVKILAVRDLIMVPRGGDIVIGLEKDSGAQRG 655
Db 121 TPGAASDRSEHDLTMDGETFSQHLQAVKILAVRDLIMVPRGGDIVIGLEKDSGAQRG 180
Oy 656 RYIAVLKARELTPHGVLDVAAYVAKDTVCTFENENTEMCLAVMRGAGAREDFIYQSYE 715
Db 181 RYIAVLKARELTPHGVLDVAAYVAKDTVCTFENENTEMCLAVMRGAGAREDFIYQSYE 240
Oy 716 ELGRLEACTRRKR 728
Db 241 ELGRLEACTRRKR 253

```

RESULT 3

OyH085 PRELIMINARY: PRT: 252 AA.

```

AC 08R085:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to hypothetical protein FLJ23119.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
Ox NCBI_TaxID=10090;
RN [1].
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RA Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC027199; AAH27199.1;
KM Hypothetical protein.
SQ SEQUENCE 252 AA; 28042 MW; 8E9BBBCAF600E95F CRC64;

```

Query Match 27.4%; Score 1046.5; DB 11; Length 252;
 Best Local Similarity 80.6%; Pred. No. 2.8e-80;
 Matches 204; Conservative 10; Mismatches 38; Indels 1; Gaps 1;

```

Oy 476 MYHSTTYOLCARFCGVPSPRLDMFPVRLDTEPPAASHTANPKYPEGDSIADVSIMYSE 535
Db 1 MYHSTTYOLCARFCGVPSPRLDMFPVRLDTEPPAASHTANPKYPEGDSIADVSIMYSE 59
Oy 536 ELGTQILIHQESLTDYCSMSYSSSPROAARSPSSLPSSPASSSVPFSTDCEDSDMLH 595
Db 61 ELGTQILIHQESLTDYCSMSYSSSPROAARSPSSLPSSPASSSVPFSTDCEDSDMLH 119
Oy 596 TPGAASDRSEHDLTMDGETFSQHLQAVKILAVRDLIMVPRGGDIVIGLEKDSGAQRG 655
Db 120 TPGAASDRSEHDLTMDGETFSQHLQAVKILAVRDLIMVPRGGDIVIGLEKDSGAQRG 179
Oy 656 RYIAVLKARELTPHGVLDVAAYVAKDTVCTFENENTEMCLAVMRGAGAREDFIYQSYE 715
Db 180 RYIAVLKARELTPHGVLDVAAYVAKDTVCTFENENTEMCLAVMRGAGAREDFIYQSYE 239
Oy 716 ELGRLEACTRRKR 728
Db 240 ELGRLEACTRRKR 252

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:19:44 ; Search time 98 seconds
(without alignments)
1330.636 Million cell updates/sec

Title: US-09-836-392-21

Perfect score: 3822

Sequence: 1 MRLRLATDAMKNFSEFRQEE.....IFVQSYELGRLACTRRKR 728

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21:*

- 1: sp._archaea:*
- 2: sp._bacteria:*
- 3: sp._fungi:*
- 4: sp._human:*
- 5: sp._invertebrate:*
- 6: sp._mammal:*
- 7: sp._mhc:*
- 8: sp._organelle:*
- 9: sp._phage:*
- 10: sp._plant:*
- 11: sp._rodent:*
- 12: sp._virus:*
- 13: sp._vertebrate:*
- 14: sp._unclassified:*
- 15: sp._virus:*
- 16: sp._bacteriap:*
- 17: sp._archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3822	100.0	1369	4	Q96JN5
2	1325	34.7	253	4	Q9H5S3
3	1046.5	27.4	252	11	Q8R085
4	1031	27.0	221	4	Q9BS11
5	402	10.5	1175	5	Q9T2M4
6	364.5	9.5	2308	5	Q9VDJ9
7	315	8.2	412	10	Q9M085
8	314	8.2	411	10	Q9Z031
9	292	7.6	1001	11	Q8VDG6
10	288.5	7.5	1020	5	Q9M313
11	288.5	7.5	1148	5	Q9SVF6
12	288.5	7.5	1161	5	Q9SUN8
13	287	7.5	416	10	Q94C42
14	278.5	7.3	546	10	Q22558
15	273.5	7.2	1066	4	Q9H2N5
16	271	7.1	411	10	Q9AWA6

17	270.5	7.1	553	10	Q81808	081808 arabidopsis
18	270.5	7.1	1171	10	Q9STG4	Q9STG4 arabidopsis
19	269.5	7.1	570	10	Q8RWL6	Q8RWL6 arabidopsis
20	268.5	7.0	438	10	Q9X187	Q9X187 arabidopsis
21	268.5	7.0	438	10	Q8S9J9	Q8S9J9 arabidopsis
22	268.5	7.0	445	10	Q9LME8	Q9LME8 arabidopsis
23	268	7.0	475	10	Q9STG5	Q9STG5 arabidopsis
24	267.5	7.0	1147	10	Q9SA26	Q9SA26 arabidopsis
25	267	7.0	1040	5	Q8T7Z0	Q8T7Z0 caenorhabd
26	266	7.0	978	5	Q8T7Z1	Q8T7Z1 caenorhabd
27	264	6.9	1553	5	Q8T219	Q8T219 dictyostell
28	263	6.9	422	5	Q23846	Q23846 dictyostell
29	263	6.9	483	10	Q8RX96	Q8RX96 arabidopsis
30	263	6.9	527	5	Q9RT25	Q9RT25 dictyostell
31	263	6.9	564	4	Q9H1Y7	Q9H1Y7 homo sapien
32	263	6.9	1029	10	Q23015	Q23015 arabidopsis
33	263	6.9	1338	5	Q23927	Q23927 dictyostell
34	262	6.9	378	10	Q9LUI6	Q9LUI6 arabidopsis
35	260.5	6.8	847	4	Q16584	Q16584 homo sapien
36	260.5	6.8	1248	10	Q9SAJ2	Q9SAJ2 arabidopsis
37	259.5	6.8	345	10	Q9SZM7	Q9SZM7 arabidopsis
38	259.5	6.8	850	11	Q9UJ15	Q9UJ15 mus musculu
39	259	6.8	525	10	Q9F1L6	Q9F1L6 arabidopsis
40	257.5	6.7	765	10	Q93YU0	Q93YU0 arabidopsis
41	255.5	6.7	763	10	Q9C903	Q9C903 arabidopsis
42	255.5	6.7	777	10	Q9C833	Q9C833 arabidopsis
43	255	6.7	525	10	Q8W0Z2	Q8W0Z2 arabidopsis
44	254	6.6	570	4	Q8W0Z2	Q8W0Z2 homo sapien
45	254	6.6	1036	4	Q8W0Z1	Q8W0Z1 homo sapien

ALIGNMENTS

RESULT 1

ID Q96JN5 PRELIMINARY; PRT; 1369 AA.

AC Q96JN5; 01-DEC-2001 (TREMBL) 19, Created)
DT 01-DEC-2001 (TREMBL) 19, Last sequence update)
DT 01-JUN-2002 (TREMBL) 21, Last annotation update)
DE KIAA1790 protein (Fragment).
GN KIAA1790.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
CC - SIMILARITY: CONTAINS 1 WD REPEAT (TRP-ASP DOMAIN).
CC EMBL: AB058693; BAB47419.1;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001005; Myb_DNA_binding.
DR Pfam: PF00069; pkinase.1.
DR ProDom: PD000001; Euk_pkinase.1.
DR PROSITE: PS00037; MYB_1; UNKNOWN.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Repeat; WD repeat.
FT NON_TER
SQ SEQUENCE 1369 AA; 153626 MW; AB058693; 88370D9 CRC64;
Query Match 100.0%; Score 3822; DB 4; Length 1369;
Best local Similarity 100.0%; Pred. No. 66-314;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

CURRENT APPLICATION NUMBER: US/10/132,382
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 7015
TYPE: DNA
ORGANISM: Homo sapiens
US-10-132-382-5

Query Match      88.7%; Score 1936.8; DB 9; Length 7015;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1

QY      1 ATGCTGAGGCACCTGCGGGCCACCGATGCCATGAAGAAGTCTCCGAGTTCGGCGAGAG 60
Db      4124 ATGCTGAGGCACCTGCGGGCCACCGATGCCATGAAGAAGTCTCCGAGTTCGGCGAGAG 4183

QY      61 GCCAGCATGCTGCACGGCGTGCAGCACCCCTGCATCTGTGGCGCTCATGTGGGATCAGCATC 120
Db      4184 GCCAGCATGCTGCACGGCGTGCAGCACCCCTGCATCTGTGGCGCTCATGTGGGATCAGCATC 4243

QY      121 CACCGGCTCTGCTGGCCCTGGAGCTCGGGCGCGCTCAGCAGAGCTCAACACCGTGTCTC 180
Db      4244 CACCGGCTCTGCTGGCCCTGGAGCTCGGGCGCGCTCAGCAGAGCTCAACACCGTGTCTC 4303

QY      181 GAGAACGCCAGAGATCTCTTATACCCCTGGGCACACATGCTCACCCAAAAAATAGCC 240
Db      4304 GAGAACGCCAGAGATCTCTTATACCCCTGGGCACACATGCTCACCCAAAAAATAGCC 4363

QY      241 TACGAGATGCGCTCGGGCGCTGCGCTTACCTGCACACAGAAAAACATCATCTTCTGTACCTG 300
Db      4364 TACGAGATGCGCTCGGGCGCTGCGCTTACCTGCACACAGAAAAACATCATCTTCTGTACCTG 4423

QY      301 AAGTGGGACAAATCTGTGTGTGTGTCCTGTGACGTCAAGAGAGACATCAACATCAAGCTTA 360
Db      4424 AAGTGGGACAAATCTGTGTGTGTGTCCTGTGACGTCAAGAGAGACATCAACATCAAGCTTA 4483

QY      361 TCTGACTACGAGGATTTGCGAGGCAAGTCAATTCATAGAGGGGCGCTTAGCGCTGAGGGCACT 420
Db      4484 TCTGACTACGAGGATTTGCGAGGCAAGTCAATTCATAGAGGGGCGCTTAGCGCTGAGGGCACT 4543

QY      421 CCTGGCTACAGGCCCCAGAGATCAGGCGCTCGCATTTGATATGATGAGAAAGTATGATG 480
Db      4544 CCTGGCTACAGGCCCCAGAGATCAGGCGCTCGCATTTGATATGATGAGAAAGTATGATG 4603

QY      481 TTCCTCTAAGGAATGTGTCTACGAGTGTGTCTACGAGGAGGAGGCGCCCTGCACTGGGGCAC 540
Db      4604 TTCCTCTAAGGAATGTGTCTACGAGTGTGTCTACGAGGAGGAGGCGCCCTGCACTGGGGCAC 4663

QY      541 CACCAAGCTCAGATTGCCAAGAAGCTGTCCAAAGGGCATCCGCCGGTTCTGGGGGACGCCG 600
Db      4664 CACCAAGCTCAGATTGCCAAGAAGCTGTCCAAAGGGCATCCGCCGGTTCTGGGGGACGCCG 4723

QY      601 GAGGAAGTGCAGTTCGCGGCACTGCAGGGCGCTCATATGATGAGTGTGGGACACTTAAGCCA 660
Db      4724 GAGGAAGTGCAGTTCGCGGCACTGCAGGGCGCTCATATGAGTGTGGGACACTTAAGCCA 4783

QY      661 GGAAGCGACCCCTGGCCCTGTGTGGGTGTGAGGCAATGAAGAACCCGATTTTTCACACC 720
Db      4784 GGAAGCGACCCCTGGCCCTGTGTGGGTGTGAGGCAATGAAGAACCCGATTTTTCACACC 4843

QY      721 TTCATGTATGAAGTGTGTGTGGGAAGCAGACAGCTTCTTTCATCCACAGGCGCAGAG 780
Db      4844 TTCATGTATGAAGTGTGTGTGGGAAGCAGACAGCTTCTTTCATCCACAGGCGCAGAG 4903

QY      781 TACACCGTGTGTGTTTGGGATGGAAGAGAGAGTCCAGGAACCTACACGCTGTGTGATACACA 840
Db      4904 TACACCGTGTGTGTTTGGGATGGAAGAGAGAGTCCAGGAACCTACACGCTGTGTGATACACA 4963

QY      841 GAGAACGGGCTCTATGAGAGGAGCAGAGATGTGCTGCCCTGGAGTGAAGGTGAGCTGTGACAG 900
Db      4964 GAGAACGGGCTCTATGAGAGGAGCAGAGATGTGCTGCCCTGGAGTGAAGGTGAGCTGTGACAG 5023

```

QY	901	CTCCAGGCTCCAGAGATCCCTGCTGGACAGCCACCGAGGACAGAAAATCTACATCTACACC	960
Db	5024	CTCCAGGCTCCAGAGATCCCTGCTGGACAGCCACCG-----	5057
QY	961	CTCAAGGGCATGTGCCCTTAAACACACCCCAACAGGCTTGGATACTCCAGCTGTCTC	1020
Db	5058	-----	5057
QY	1021	ACCTGCTTCTTGCCGCGTGCCTTTATTAAGAAATTCCTACCTGCTGTAGCGGCGCTC	1080
Db	5058	-----AGAAATCTTACTCTGTGTAGCGGCGCTC	5086
QY	1081	GCCGATGGGCTTGTGGCTGTGTTCCTCCGCTGTGTCGGGGGACCCCAAGAGACGTGCTCC	1140
Db	5087	GCCGATGGGCTTGTGGCTGTGTTCCTCCGCTGTGTCGGGGGACCCCAAGAGACGTGCTCC	5146
QY	1141	TACCTGTGCTCACACACAGCCCAACAGGCTCCAAAGTTGAGCATGCGGGATGAAGACGACGG	1200
Db	5147	TACCTGTGCTCACACAGCCCAACAGGCTCCAAAGTTGAGCATGCGGGATGAAGACGACGG	5206
QY	1201	CAGAACCCCTACCCAGTAGAAGAGCCATGAGAGTGTTCACACGCGGCTGTAGGCTGTGTAC	1260
Db	5207	CAGAACCCCTACCCAGTAGAAGAGCCATGAGAGTGTTCACACGCGGCTGTAGGCTGTGTAC	5266
QY	1261	AGCAATGGGCGCGGCGCTCTTGTTCATCGACTGTGCTCCCTGGAGATGTGACGGCGGCTG	1320
Db	5267	AGCAATGGGCGCGGCGCTCTTGTTCATCGACTGTGCTCCCTGGAGATGTGACGGCGGCTG	5326
QY	1321	GAGCCCTACATGGCCCCCTCCATGTGTAGTCAGTGTGTGACGTCTGAGGGCAGAGGG	1380
Db	5327	GAGCCCTACATGAGCCCCCTCCATGTGTAGTCAGTGTGTGACGTCTGAGGGCAGAGGG	5386
QY	1381	GAGGAGGTGCTGTGCTGCTCGATGATGACAAAGGCAACTCTTGATGTATGACACTCCACC	1440
Db	5387	GAGGAGGTGCTGTGCTGCTCGATGATGACAAAGGCAACTCTTGATGTATGACACTCCACC	5446
QY	1441	ACCTACACAGCTGTGCTGCCGCTACTTCTGGGGGGTCCCAAGCCCCCTCAGGACATGTT	1500
Db	5447	ACCTACACAGCTGTGCTGCCGCTACTTCTGGGGGGTCCCAAGCCCCCTCAGGACATGTT	5506
QY	1501	CCCGTGCGGCCCTTGGACACAGGAACCCCGGGCAGCCACACAGGCGCAACCCAAAGTG	1560
Db	5507	CCCGTGCGGCCCTTGGACACAGGAACCCCGGGCAGCCACACAGGCGCAACCCAAAGTG	5566
QY	1561	CCTGAGGGGGACCTCATTCGGGGAGCTGAGATATATACAGTAGAGAGCTGTGGCAGCAG	1620
Db	5567	CCTGAGGGGGACCTCATTCGGGGAGCTGAGATATATACAGTAGAGAGCTGTGGCAGCAG	5628
QY	1621	ATTCGATCCACAGGATACACTCTAGTACTGTCTCATGTCTGTCTACTCTCATTC	1680
Db	5627	ATTCGATCCACAGGATACACTCTAGTACTGTCTCATGTCTGTCTACTCTCATTC	5686
QY	1681	CCACCCCGCCAGGCTGCCAGGTGCCCTCAAAGCTTCCCAAGCTCCCGACAGATTCCTCC	1740
Db	5687	CCACCCCGCCAGGCTGCCAGGTGCCCTCAAAGCTTCCCAAGCTCCCGACAGATTCCTCC	5746
QY	1741	AGTGTGCTTCTCCACAGCATGTGGGAGATCAGATGATGATACAGGCGCGGCTGTGCC	1800
Db	5747	AGTGTGCTTCTCCACAGCATGTGGGAGATCAGATGATGATACAGGCGCGGCTGTGCC	5806
QY	1801	TCCGACAGGCTTGAGCATGTGACTACCCCATGAGAGGAGGAGACTTTCAGTCAGTCAGC	1860
Db	5807	TCCGACAGGCTTGAGCATGTGACTACCCCATGAGAGGAGGAGACTTTCAGTCAGTCAGC	5866
QY	1861	CAGGCGCTGAGATCTGTGCGCTGAGAGACCTAATTGGGTCCCAAGCGCGGCTGTGAAT	1920
Db	5867	CAGGCGCTGAGATCTGTGCGCTGAGAGACCTAATTGGGTCCCAAGCGCGGCTGTGAAT	5926
QY	1921	GTTATGCTATGATGCTGGGAGAGATTCGGAAGCCGAGGCGGCGGCGGAGTATTCGCTC	1980
Db	5927	GTTATGCTATGATGCTGGGAGAGATTCGGAAGCCGAGGCGGCGGAGTATTCGCTC	5986


```

Oy 1981 TTAAGAGCCGAGAGCTGATCCGACATGGGGTCTGTGGATGCTGCGTGGTGGCAAG 2040
Db 5987 TTTAAAGCCGAGAGCTGATCTCCGCAATGGGGTCTGTGGATGCTGCGTGGTGGCAAG 6046
Oy 2041 GACACTGTGTGTGTCACACTTTGAAAATGAAAACACAGAGTGGTGGCTGGCTGTGAGAG 2100
Db 6047 GACACTGTGTGTGTCACACTTTGAAAATGAAAACACAGAGTGGTGGCTGGCTGTGAGAG 6106
Oy 2101 GCGTGGGGGCGCAGGAGTTCGACATTTCTACAGTCCCTACGAGAGAGCTGGGCGCGCTG 2160
Db 6107 GCGTGGGGGCGCAGGAGTTCGACATTTCTACAGTCCCTACGAGAGAGCTGGGCGCGCTG 6166
Oy 2161 GAGGCTTGCACTGCGACAGAGAAG 2184
Db 6167 GAGGCTTGCACTGCGACAGAGAAG 6190

RESULT 3
US-10-132-382-1
; Sequence 1, Application US/1012382
; Publication NO. US2003004569A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-382-1

```

Query Match	88.7%	Score 1936.8	DB 9	Length 7093
Best Local Similarity	94.6%	Pred. No. 0		
Matches 2065	Conservative	0	Mismatches 2	Indels 117
				Gaps 1
QY 1	ATGCTGAGGCACTGCGGGCCACGCATGCGCATGAAGAATTCCTCCGAGTCCGGCAGAG	60		
Db 4202	ATGCTGAGGCACTGCGGGCCACGCATGCGCATGAAGAATTCCTCCGAGTCCGGCAGAG	4261		
QY 61	GCCAGCATGCTGCACGCGCTGCAGACCCCTGCATCGTGGCGCTCATCGGCATCAACATC	120		
Db 4262	GCCAGCATGCTGCACGCGCTGCAGACCCCTGCATCGTGGCGCTCATCGGCATCAACATC	4321		
QY 121	CACCGGCTCTGCTTGCCCTGGAGCTCGCGCCGCTCAGCAGCCTCAACACCGTGTGCC	180		
Db 4332	CACCGGCTCTGCTTGCCCTGGAGCTCGCGCCGCTCAGCAGCCTCAACACCGTGTGCC	4381		
QY 181	GAGAACGGCAGAGATTCCTTCCTTATACCCCTGGGACACATCTCAACCCAAAAATAGCC	240		
Db 4382	GAGAACGGCAGAGATTCCTTCCTTATACCCCTGGGACACATCTCAACCCAAAAATAGCC	4441		
QY 241	TACCATGTCGCTCGGGGCTTGCTACCGTCACAGCAAGAAAAATCATCTTCGTGACCTG	300		
Db 4442	TACCATGTCGCTCGGGGCTTGCTACCGTCACAGCAAGAAAAATCATCTTCGTGACCTG	4501		
QY 301	AAGTGGACAACATTCGTGTGTGTGCCCTTGACGTCAAGGAGACATCAACATCAAGCTA	360		
Db 4502	AAGTGGACAACATTCGTGTGTGTGCCCTTGACGTCAAGGAGACATCAACATCAAGCTA	4561		
QY 361	TCTGACTACGGGATTTCGAGGCAAGTCATTCATGTAGAGGGGCCCTAGGCGCTGGAGGGCACT	420		
Db 4562	TCTGACTACGGGATTTCGAGGCAAGTCATTCATGTAGAGGGGCCCTAGGCGCTGGAGGGCACT	4621		
QY 421	CCTGCTACCAAGCCCCCAGAGATCAAGGCTCGCATTTATATGANTGAGAAGGTAGATATG	480		
Db 4622	CCTGCTACCAAGCCCCCAGAGATCAAGGCTCGCATTTATATGANTGAGAAGGTAGATATG	4681		
QY 481	TTTCTCTATGAAATGTGCTCTACAGAGTTGCTGTAGAGACAGCGCCCTCACTGGGCCAC	540		

Db	4682	TTTCCTTATGGAATGATGTCCTTACGAATTCCTCTCAGAGACAGCGCCCTCACTGGGCCAC	4741
Qy	541	CACGACGCCGATTTGCCAGAAAGACTGCTCCAAAGGATCCGCGCGGTTCTGGGGCAGCG	600
Db	4742	CACGAGCTCCAGATTTGGCAGAAAGCTGTCCAAAGGATCCGCGCGGTTCTGGGGCAGCG	4801
Qy	601	GAGGAAGTGCAGTTCCGGCGACTGACGGCGCTCATGTGAGTGTGGACACTAAGCCA	660
Db	4802	GAGGAAGTGCAGTTCCGGCGACTGACGGCGCTCATGTGAGTGTGGACACTAAGCCA	4861
Qy	661	GAGGAAGCAGCCGTCGGCCGTGGTGGTGGTGAAGCAATGAAGACCCGACTTTGGCAC	720
Db	4862	GAGGAAGCAGCCGTCGGCCGTGGTGGTGGTGAAGCAATGAAGACCCGACTTTGGCAC	4921
Qy	721	TTTCATGTATGAATGCTGCTGTGGGAACACACAGCCTTTCTTCATCCAGGGCCAGAG	780
Db	4922	TTTCATGTATGAATGCTGCTGTGGGAACACACAGCCTTTCTTCATCCAGGGCCAGAG	4981
Qy	781	TACACCGTGTGTTTTGGATGGAAAAAGAGAGTCCAGAACTACACGGTGGTGAACA	840
Db	4982	TACACCGTGTGTTTTGGATGGAAAAAGAGAGTCCAGAACTACACGGTGGTGAACA	5041
Qy	841	GAGAAAGGCCCTCATGGAGAGGACAGAGATGTGTGGCCCTGGGATGAAGGTGAGCTGGC	900
Db	5042	GAGAAAGGCCCTCATGGAGAGGACAGAGATGTGTGGCCCTGGGATGAAGGTGAGCTGGC	5101
Qy	901	CTCCAGGTCAGAGATCCCTGTGGACAGCCAGCAGAGACCAGAAATCTACATCTACAC	960
Db	5102	CTCCAGGTCAGAGATCCCTGTGGACAGCCAGCAGAGACCAGAAATCTACATCTACAC	5135
Qy	961	CTCAAGGCGATGTGCCCTTAACACACCCCAACAGCCCTTGATATCTCAGCTGTCTC	1020
Db	5136	-----	5135
Qy	1021	ACCTGCTTCTGGGCGGTGCGCTGTTATTAAGAAATTCCTACCTGGTCTTACGGGCGCTC	1080
Db	5136	-----AATAATTCCTACCTGGTCTTACGGGCGCTC	5166
Qy	1081	GCCGATGGGCTTGTGGCTGTGTTTCCCGTGGTGCAGGGGACCCCAAGAGACAGCTGCTC	1140
Db	5165	GCCGATGGGCTTGTGGCTGTGTTTCCCGTGGTGCAGGGGACCCCAAGAGACAGCTGCTC	5222
Qy	1141	TACCTGTCTTACACACACAGCCACAGTCCCAATTCAGCATGCGGATGAGAGCGACGG	1200
Db	5225	TACCTGTCTTACACACAGCCACAGTCCCAATTCAGCATGCGGATGAGAGCGACGG	5288
Qy	1201	CAGAACCCCTTCCAGAGGAAGGCGCATGAGAGTGGTAAACAGGGGCTCTAGGTCGTGATC	1260
Db	5285	CAGAACCCCTTCCAGAGGAAGGCGCATGAGAGTGGTAAACAGGGGCTCTAGGTCGTGATC	5344
Qy	1261	AGCAATGGGCGGGCGCTCTGTTCATGATGTGTGCTCCCTGGAGATCTGCAGGGCGCTG	1320
Db	5345	AGCAATGGGCGGGCGCTCTGTTCATGATGTGTGCTCCCTGGAGATCTGCAGGGCGCTG	5404
Qy	1321	GAGCCCTTACATGGCCCCCTCCATGTTAGTCAATGCTGTGCAAGCTCTGAGGGCAGAGG	1380
Db	5405	GAGCCCTTACATGGCCCCCTCCATGTTAGTCAATGCTGTGCAAGCTCTGAGGGCAGAGG	5464
Qy	1381	GAGGAGGTCGTGTGCTGCTGCGTGAATGACAAAGCCCAACTCTTGGTATGTACACTGCAC	1440
Db	5465	GAGGAGGTCGTGTGCTGCTGCGTGAATGACAAAGCCCAACTCTTGGTATGTACACTGCAC	5524
Qy	1441	ACCTACACAGCTGTGGCCGCGATCACTTGTGGGGGTCCCGAGCCCTCTAGGGACATGTT	1500
Db	5525	ACCTACACAGCTGTGGCCGCGATCACTTGTGGGGGTCCCGAGCCCTCTAGGGACATGTT	5588
Qy	1501	CCCGTGGGCCCTTGGACAGGAACCCCGGCGACGACACACGGCCCAACCAAGAGTG	1560
Db	5585	CCCGTGGGCCCTTGGACAGGAACCCCGGCGACGACACACGGCCCAACCAAGAGTG	5644
Qy	1561	CCTGAGGGGACTCTCATGCGGAGCGTAGCATATTTACAGTGAAGAGTGGGCACGAG	1620
Db	5645	CCTGAGGGGACTCTCATGCGGAGCGTAGCATATTTACAGTGAAGAGTGGGCACGAG	5704

QY	1821	ATTGTATTCACAGGAATACACTGACTGACTGCTGCATGCTCCATCCTCATC	1680
Db	5705	ATCTCGTATCCACCAGGAATACACTGACTGACTGCTGCATCCTCATCCTCATC	5764
QY	1681	CCAACCCGCGAGGCTGCCAGTGTCGCCCTCAACCTCCAGCTCCCGACAAGTTCTTC	1740
Db	5765	CCAACCCGCGCGAGGCTGCCAGTGTCGCCCTCAACCTCCAGCTCCCGACAAGTTCTTC	5824
QY	1741	AGTGTGCTTTTCACACCGACTGCGAGGACTCAGACATGCTAATACGCCCGGTGCTGC	1800
Db	5825	AGTGTGCTTTTCACACCGACTGCGAGGACTCAGACATGCTAATACGCCCGGTGCTGC	5884
QY	1801	TCCGACACAGTGTCTAGACATACCTGCACCCCACATGAGCGGGGAGAACTTTAGCACAGACTG	1860
Db	5885	TCCGACACAGTGTCTAGACATACCTGCACCCCACATGAGCGGGGAGAACTTTAGCACAGACTG	5944
QY	1861	CAGGCCCTGTAAGAATCCGCGCCGTCAAGACACTCATTTGGGTCCCCAGCGCGGTGGAGAT	1920
Db	5945	CAGGCCCTGTAAGAATCCGCGCCGTCAAGACACTCATTTGGGTCCCCAGCGCGGTGGAGAT	6004
QY	1921	GTTATCTGTCAATTGGCCGTGAGAAAGGATTTCTGAAGCCCAAGGGGGCCGAGTCATTGCCGTC	1980
Db	6005	GTTATCTGTCAATTGGCCGTGAGAAAGGATTTCTGGGCCCAAGGGGGCCGAGTCATTGCCGTC	6064
QY	1981	TTTAAACACCCGAGAGCTGCATCCCGCATGGGGGTGTGTGTGATGTGTCGCCGTGTGGCAAG	2040
Db	6065	TTTAAACACCCGAGAGCTGCATCCCGCATGGGGGTGTGTGTGATGTGTCGCCGTGTGGCAAG	6124
QY	2041	GACACTGTTGTGTGCACCTTTGTAAAAATGAATAACACAGATGTGTGCTGTGCTGTGAGG	2100
Db	6125	GACACTGTTGTGTGCACCTTTGTAAAAATGAATAACACAGATGTGTGCTGTGCTGTGAGG	6184
QY	2101	GCGTGGGGCGCCAGGAGGTTGCACATTTGTATACAGTCTCTACAGAGAGCTGGGCGGCTG	2160
Db	6185	GCGTGGGGCGCCAGGAGGATTGCACATTTGTATACAGTCTCTACAGAGAGAGCTGGGCGGCTG	6244
QY	2161	GAGGCTTGCACTGCGCAAGAGAAG 2184	
Db	6245	GAGGCTTGCACTGCGCAAGAGAAG 6268	

```

RESULT 4
US-10-132-382-7
; Sequence 7, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 7229
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-382-7

Query Match      88.7%; Score 1936.8; DB 9; Length 7229;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1;

QY      1  ATGCTGAGCGACCTGCGGGCCACCAGATGGCATGAGAAATTTCTCCGAGTTCGGGAGAG 60
        |||||||
DB      4338  ATGCTGAGCGACCTGCGGGCCACCAGATGGCATGAGAAATTTCTCCGAGTTCGGGAGAG 4397

QY      61  GCCAGATCTCTACAGCGCTGCAGACACCCTCATCTGTTGGCGCTCATCGCATCAGCATC 120
        |||||||
DB      4398  GCCAGATCTCTACAGCGCTGCAGACACCCTCATCTGTTGGCGCTCATCGCATCAGCATC 4457

QY      121  CACCGCTCTGCTTGGCCCTTGAGACTGCGCGCGCTCAGCAGCCTCAGCACCGTGTCTGTC 180

```

Db	4458	CACCCGTCGTGCTTGCCCTCGGAGCTCGGGCGCGCTCAGACAGCTCAACACCGTGTGTCC	4517
QY	181	GAGAAAGCCAGAGATTCTTCTTTATACCCCTGGGACATGCTCACCCAAAAATAGCC	240
Db	4518	GAGAAAGCCAGAGATTCTTCTTTATACCCCTGGGACATGCTCACCCAAAAATAGCC	4577
QY	241	TACAGATGCGCCCGGGGCGTACCGCGCAAGAAAAATCATCTTGTGACCTG	300
Db	4578	TACAGATGCGCCCGGGGCGTACCGCGCAAGAAAAATCATCTTGTGACCTG	4537
QY	301	AAGTCGACAACTTCTGGTGTGTCCCTTGTGACGTCGAAGAGCATCATCAAGCTA	360
Db	4638	AAGTCGACAACTTCTGGTGTGTCCCTTGTGACGTCGAAGAGCATCATCAAGCTA	4697
QY	361	TCTGACTACGGGATTTTCGAGGAGTCAATTCATGAGGGCGCCTTAGCGTGGAGCACT	420
Db	4698	TCTGACTACGGGATTTTCGAGGAGTCAATTCATGAGGGCGCCTTAGCGTGGAGCACT	4757
QY	421	CTTGCGCTACCAAGGCCCCCAAGAGATCAAGGCGCTGCATTTGATATGATGAAGGTGATATG	480
Db	4758	CTTGCGCTACCAAGGCCCCCAAGAGATCAAGGCGCTGCATTTGATATGATGAAGGTGATATG	4817
QY	481	TTTCCTATGAATGATGCTCTACAGTGTGTGTGAGACAGCGCCCTGCACCTGCGCAC	540
Db	4818	TTTCCTATGAATGATGCTCTACAGTGTGTGTGAGACAGCGCCCTGCACCTGCGCAC	4877
QY	541	CACGAGCTCAGATTGCGCAAGAACTGTCCAAAGGCAATCCGCCGGTTCTGGGCGACCG	600
Db	4878	CACGAGCTCAGATTGCGCAAGAACTGTCCAAAGGCAATCCGCCGGTTCTGGGCGACCG	4937
QY	601	GAGAAATGCAATTCGCGGCGCATGCGAGGGGCTCATGATGATGATGGGAACTAATACCA	660
Db	4938	GAGAAATGCAATTCGCGGCGCATGCGAGGGGCTCATGATGATGATGGGAACTAATACCA	4997
QY	661	GAGAAAGCAAGCGCTGCGGCTGTGATGATGAGCCAGATGAAGAACCCACTTTGGCAC	720
Db	4998	GAGAAAGCAAGCGCTGCGGCTGTGATGATGAGCCAGATGAAGAACCCACTTTGGCAC	5057
QY	721	TTTCATGATGAATGCTGTGTGGAAAGCAAGACGCTTCTCATCCAGGGCCAGAG	780
Db	5058	TTTCATGATGAATGCTGTGTGGAAAGCAAGACGCTTCTCATCCAGGGCCAGAG	5117
QY	781	TACACCGTGTGTTTGGATGTGAAAAAGAGAGTCCAGAACTACAGGTGTGTAACA	840
Db	5118	TACACCGTGTGTTTGGATGTGAAAAAGAGAGTCCAGAACTACAGGTGTGTAACA	5177
QY	841	GAGAAAGGCTCATGAGAGTGCAGAGAGATGTGCTGCGGTGAAGATGAAGTGTGCCAG	900
Db	5178	GAGAAAGGCTCATGAGAGTGCAGAGAGATGTGCTGCGGTGAAGATGAAGTGTGCCAG	5237
QY	901	CTTCAGGTCAGAGATCCCTGTGGAGACGACCCAGAGACAGAAAAATCTACATCTACAC	960
Db	5238	CTTCAGGTCAGAGATCCCTGTGGAGACGACCCAGAGACAGAAAAATCTACATCTACAC	5271
QY	961	CTCAAGGCGATGTGCCCCCTTAAACACACCCCAAGGCTTGGATATCCAGCTGTGCTC	1020
Db	5272	CTCAAGGCGATGTGCCCCCTTAAACACACCCCAAGGCTTGGATATCCAGCTGTGCTC	5271
QY	1021	ACCTGCTTCTTGCGGCGTGCCTGTTATTTAAAAAGATTTCTACCTGATGCTTAGCGGAGCTC	1080
Db	5272	ACCTGCTTCTTGCGGCGTGCCTGTTATTTAAAAAGATTTCTACCTGATGCTTAGCGGAGCTC	5300
QY	1081	GCCGATGGGCTTGTGGCTGTGTTTCCCGTGTGCGGGGACCCCAAGAGACAGCTGTCC	1140
Db	5301	GCCGATGGGCTTGTGGCTGTGTTTCCCGTGTGCGGGGACCCCAAGAGACAGCTGTCC	5360
QY	1141	TACCTGTCTCACAACAAGCCAAAGTCCAAATTTAGCATGCGGATGTAAGACGACAGG	1200
Db	5361	TACCTGTCTCACAACAAGCCAAAGTCCAAATTTAGCATGCGGATGTAAGACGACAGG	5420
QY	1201	GAGAACCCCTCCAGTAAGAGGCAATGAGATGTGTACAGGGGCTGTAGGTCTGTGATC	1260

```

Db 5421 CAGAACCCCTACCCAGTGAAGCCATGAGAGTGTCTCAACAGCGGCTCTGAGTGTGTAC 5480
Qy 1261 AGCAATGGGCGGCGCTCTCTTGCATGACGTGTGCTCCCTGGAGATTCGACAGGGGCTG 1320
Db 5481 AGCAATGGGCGGCGCTCTCTTGCATGACGTGTGCTCCCTGGAGATTCGACAGGGGCTG 5540
Qy 1321 GAGCCCTACATGAGCCCTCTCATGTTAGTACGTGCTGTCAGCTCTGAGGGCAGAGG 1380
Db 5541 GAGCCCTACATGAGCCCTCTCATGTTAGTACGTGCTGTCAGCTCTGAGGGCAGAGG 5600
Qy 1381 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 5601 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5660
Qy 1441 ACCTACACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 5661 ACCTACACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5720
Qy 1501 CCCGTGCGGCGCTTGGACACGGAACCCCGGCGAGCCAGCCACAGGCCAACCCAAAGCTG 1560
Db 5721 CCCGTGCGGCGCTTGGACACGGAACCCCGGCGAGCCAGCCACAGGCCAACCCAAAGCTG 5780
Qy 1561 CCTGAGGGGGATCTCCATTCGCGGACGTGACATCATATACAGTGAAGAGTGGGCGACG 1620
Db 5781 CCTGAGGGGGATCTCCATTCGCGGACGTGACATCATATACAGTGAAGAGTGGGCGACG 5840
Qy 1621 ATCTGATGACACAGGATCACTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 5841 ATCTGATGACACAGGATCACTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5900
Qy 1681 CCACCCCGGCGAGGCTGCGAGGTCCCGCTCAAGCCTCCCGAGTCCCGACCAAGTTCTTCC 1740
Db 5901 CCACCCCGGCGAGGCTGCGAGGTCCCGCTCAAGCCTCCCGAGTCCCGACCAAGTTCTTCC 5960
Qy 1741 AATGTGCTTCTTCTCCACGACGTGCGAGGACTCAGACATCTACATACGCCCGGTGCTGCC 1800
Db 5961 AATGTGCTTCTTCTCCACGACGTGCGAGGACTCAGACATCTACATACGCCCGGTGCTGCC 6020
Qy 1801 TCCGACAGGTCTGAGCATGACCTGACCCCATGAGAGGGGAGAGCCCTTCAGCCAGACCTG 1860
Db 6021 TCCGACAGGTCTGAGCATGACCTGACCCCATGAGAGGGGAGAGCCCTTCAGCCAGACCTG 6080
Qy 1861 CAGGCGGTGAGATCTCTGCGCGCTCAGAGACCTCATTTGGTCCCGCGGCGGTGAGAT 1920
Db 6081 CAGGCGGTGAGATCTCTGCGCGCTCAGAGACCTCATTTGGTCCCGCGGCGGTGAGAT 6140
Qy 1921 GTTATGCTATTTGGCTTGAGAAAGATTGTGAAGCCAGCGGCGGTGAGAT 1980
Db 6141 GTTATGCTATTTGGCTTGAGAAAGATTGTGAAGCCAGCGGCGGTGAGAT 6200
Qy 1981 TTTAAAGCCGAGAGTGTGCTGCGATGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
Db 6201 TTTAAAGCCGAGAGTGTGCTGCGATGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6260
Qy 2041 GACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
Db 6261 GACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6320
Qy 2101 GGCTGGGGGGCGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
Db 6321 GGCTGGGGGGCGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6380
Qy 2161 GAGGCTGTGACTGTGCAAGAGG 2184
Db 6381 GAGGCTGTGACTGTGCAAGAGG 6404

```

```

RESULT 5
US-10-132-382-3
; Sequence 3, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM

```

```

; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 7307
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-382-3

Query Match      88.7%; Score 1936.8; DB 9; Length 7307;
Best Local Similarity 94.6%; Pred. No. 0; Mismatches 2; Indels 117; Gaps 1;
Matches 2065; Conservative 0;

Qy 1 ATGCTGAGGACCTGCGGGCCACCGATGCAATGAAAGAACTTCTCCGAGTCCGCGAGAG 60
Db 4416 ATGCTGAGGACCTGCGGGCCACCGATGCAATGAAAGAACTTCTCCGAGTCCGCGAGAG 4475
Qy 61 GCCAGATCTCTCAGCGGCTGCGACACCCCTCATCTGTGGCGCTCATCGGATCAGCATC 120
Db 4476 GCCAGATCTCTCAGCGGCTGCGACACCCCTCATCTGTGGCGCTCATCGGATCAGCATC 4535
Qy 121 CACCGGCTGCTGCGGCTGAGGCTGCGGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 180
Db 4536 CACCGGCTGCTGCGGCTGAGGCTGCGGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 4595
Qy 181 GAGAACGCCAGAGATCTTCTTATACCCCTGGAGACATGCTCACCCAAAAATAGCC 240
Db 4596 GAGAACGCCAGAGATCTTCTTATACCCCTGGAGACATGCTCACCCAAAAATAGCC 4655
Qy 241 TACAGATGCTGCGGCGCTGCGCTTACTGCAAGAAAGAAATCATCTTCTGTGACTG 300
Db 4655 TACAGATGCTGCGGCGCTGCGCTTACTGCAAGAAAGAAATCATCTTCTGTGACTG 4715
Qy 301 AAGTGGAGACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db 4716 AAGTGGAGACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4775
Qy 361 TCTGACTAGGAGATTTGAGGAGTATTCATTCATGAGGGGCGGCTGAGGGGTGAGGGGACT 420
Db 4776 TCTGACTAGGAGATTTGAGGAGTATTCATTCATGAGGGGCGGCTGAGGGGTGAGGGGACT 4835
Qy 421 CCTGGCTACAGAGCCCGCAGAGATCAGGGCTGCTGATTTATGATGAGAGTATATG 480
Db 4836 CCTGGCTACAGAGCCCGCAGAGATCAGGGCTGCTGATTTATGATGAGAGTATATG 4895
Qy 481 TTTCTCTATGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db 4896 TTTCTCTATGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4955
Qy 541 CACGAGCTCCAGATTTGCCAAGAGTGTGCCAAGGGGATCCGCGGCTTGTGGGCGAGCG 600
Db 4956 CACGAGCTCCAGATTTGCCAAGAGTGTGCCAAGGGGATCCGCGGCTTGTGGGCGAGCG 5015
Qy 601 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db 5016 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5075
Qy 661 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db 5076 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5135
Qy 721 TTTATGTATGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 5136 TTTATGTATGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5195
Qy 781 TACACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 5196 TACACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5255
Qy 841 GAGAGGGGCTCATGTGAGAGTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900

```


RESULT 7
US-09-935-464-2
Sequence 2, Application US/09935464
Publication No. US20030027153A1
GENERAL INFORMATION:
APPLICANT: Meyer, Joanne
APPLICANT: Barrington-Martin, Rory
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
FILE REFERENCE: 3322/1H702 US1
CURRENT FILING DATE: 2001-08-23
PRIOR APPLICATION NUMBER: US 09/757,300
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1383
TYPE: DNA
ORGANISM: Homo sapiens
US-09-935-464-2

Query Match 2.5%, Score 54.2; DB 9; Length 1383;
Best Local Similarity 49.9%; Pred. No. 1.1e-05;
Matches 168; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

OY 240 CTACAGATCGCGCTGGGCTGAGCTTACCTGACAAAGAAACATCATCTTGTGACCT 299
DB 372 CCAGAGAGTCTTGTGCGAGTGAATACCTACATGAGATGCGATCCACAGAGACT 431
OY 300 GAAGTCGACACATCTTGTGTGCTGCTGACCTGACGCTCAAGAGACATCATCAAGCT 359
DB 432 AAAGCCGAAACCTGCT-----TTACCTTACCCCTGAGAGACTCATACATCATGAT 485
OY 360 ATCTACTACAGGATTTGAGGCACTCATTCATGAGGCGCCCTGAGCGCTGAGAGGAC 419
DB 486 CACTGACTTGTGCTGCTCAAGATGGAACAGATGACATGCTCAGCTGTGGGAC 545
OY 420 TCTGCTACAGGCGCCGACAGATGAGGCTGCTGATGATGATGAGAGATGATAT 479
DB 546 CCCAGGCTACGCTGCTCCAGAGTCTGCGCCGAGAAACCTTACAGAGAGCTGTGATG 605
OY 480 GTTCTCTATGGAATGTGCTCTACAGTGTGCTGACAGAGCGCCCTGACACTGAGGCA 539
DB 606 CTGCTCATCGGCGCTCATACATCATATGCTGTGATGATCCGCCATTTATGAAGA 665
OY 540 CCACAGCTCCAGATTTGCCAAGAGCTTCCAGGCG 576
DB 666 AACGAGTCTAAGCTTTTCAGAAAGATCAAGAGGCG 702

RESULT 8
US-09-935-464-4
Sequence 4, Application US/09935464
Publication No. US20030027153A1
GENERAL INFORMATION:
APPLICANT: Meyer, Joanne
APPLICANT: Barrington-Martin, Rory
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
FILE REFERENCE: 3322/1H702 US1
CURRENT FILING DATE: 2001-08-23
PRIOR APPLICATION NUMBER: US 09/757,300
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 1738
TYPE: DNA
ORGANISM: Homo sapiens

US-09-935-464-4

Query Match 2.5%; Score 54.2; DB 9; Length 1738;
Best Local Similarity 49.9%; Pred. No. 1.2e-05;
Matches 168; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

OY 240 CTACAGATCGCGCTGGGCTGAGCTTACCTGACAAAGAAACATCATCTTGTGACCT 299
DB 369 CCAGAGAGTCTTGTGCGAGTGAATACCTACATGAGATGCGATCCACAGAGACTT 448
OY 300 GAAGTCGACACATCTTGTGTGCTGCTGACCTGACGCTCAAGAGACATCATCAAGCT 359
DB 449 AAAGCCGAAACCTGCT-----TTACCTTACCCCTGAGAGACTCATACATCATGAT 502
OY 360 ATCTACTACAGGATTTGAGGCACTCATTCATGAGGCGCCCTGAGCGCTGAGAGGAC 419
DB 503 CACTGACTTGTGCTGCTCAAGATGGAACAGATGACATGATCTCCACTGCTGTGGGAC 562
OY 420 TCTGCTACAGGCGCCGACAGATGAGGCTGCTGATGATGATGAGAGATGATAT 479
DB 563 CCCAGGCTACGCTGCTCCAGAGTCTGCGCCGAGAAACCTTACAGAGAGCTGTGATG 622
OY 480 GTTCTCTATGGAATGTGCTCTACAGTGTGCTGACAGAGCGCCCTGACACTGAGGCA 539
DB 623 CTGCTCATCGGCGCTCATACATCATATGCTGTGATGATCCGCCATTTATGAAGA 682
OY 540 CCACAGCTCCAGATTTGCCAAGAGCTTCCAGGCG 576
DB 683 AACGAGTCTAAGCTTTTCAGAAAGATCAAGAGGCG 719

RESULT 9
US-10-096-960-1
Sequence 1, Application US/10096960
Patent No. US2002013235A1
GENERAL INFORMATION:
APPLICANT: YE, Jane et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C1001158DIV
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 09/800,960
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2061-
TYPE: DNA
ORGANISM: Homo sapien
US-10-096-960-1

Query Match 2.4%; Score 51.4; DB 12; Length 2061;
Best Local Similarity 47.4%; Pred. No. 8e-05;
Matches 154; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

OY 242 ACCAGATGCGCTGGGCTGAGCTTACCTGACAAAGAAACATCATCTTGTGACCTGA 301
DB 465 ATCAATTTCTGAGAGTGTTAACCATCCACACAGATGATGCTCCACAGGAGCTGA 524
OY 302 AGTCGACACATCTGCTGTGCTGCTTGAAGTCAAGAGACATCAACATCAAGCTAT 361
DB 525 AGCTGAGAACTGCTGCTGCGAGTAATATCAAGAGTCCGCTCAAGCTGCTGATTT 584
OY 362 CTGACTACGGGATTTGAGGAGTCAATTCATGAGGCGCCCTGAGCGCTGAGGCACTC 421
DB 585 TTGGCTAGCCATGAGATACAGGAGAGAGAGAGAGGCTGTGTTGGTTTCTGTCGACACC 644
OY 422 CTGCTACAGGCGCCGACAGATCAGGCTGCTGATTTGATGATGAGAGATGATATGT 481
DB 645 CAGTTACTTGTCCCTGAGGCTTTGAGAGAAAGATCCCTATGAGAAACCTGTGATATCT 704

OY 349 AACATCAAGCTATCTGACTACGGGATTTGAGGCACTATTCATGAGGCGCCCTAGGC 408
DB 466 GATTTCGGCATGCGCCGGCCATTCGCCAGCGGCAACAGCGTATGCCAGCCAGCA 525
OY 409 GTGAGGCGCATCCCTGGGTACCAAGCCCAAGATCAGAGCCCTCGATGTATATATAG 468
DB 526 GTGATCGGACAGCGCCAGTACCTGTACCCGAAACAGGCCCCGGGGATTCCTCAGCGC 585
OY 469 AAGGTAGATATCTCTATGTAATGGTCTCTACGATTTCTCTACAGAGCGCCCT 528
DB 586 CGATCCGATGTATCTCTGTGGCTGTCTTTATGAAGTCTCACCAGGAGCCACT 645
OY 529 GCATGCGGCGCACACAGCTCCAGATTCGCAAGAGCTGTCCAGGCGATCCCGGTT 588
DB 646 TTCACCGGGGACATCACCCTGCTGGGTCTACCAACATGTGCGGAGAGCCGATCCCA 705
OY 589 CTGGGCGACCGGAGAGAGTTCGCGGCGACTGCGAGCCCTCATGATGAGTCTGG 648
DB 706 CTTGCGGCGGCGAGAGGCTCTCCGCGGACCTGAGACGCGTCTCTCAAGGCGCTG 765
OY 649 GACACTAAGCGAGAGGACCGGCTGTGCGGTGTGAGC 693
DB 766 GCCAAAATCCGAAAAACCTATATCAGACGCGGAGATCGC 810

RESULT 15

US-09-823-038A-58
; Sequence 58, Application US/09823038A
; Patent No. US2002058335A1
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murlison, Greg
; TITLE OF INVENTION: Compositions Isolated From Stromal Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1037c3
; CURRENT APPLICATION NUMBER: US/09/823, 038A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 3503
; TYPE: DNA
; ORGANISM: Mouse
US-09-823-038A-58

Query Match 2.2%; Score 48.2; DB 10; Length 3503;

Best Local Similarity 51.7%; Pred. No. 0.00078; Mismatches 143; Indels 12; Gaps 2;

OY 210 CCTGGACACATGCTCCACCAAAAATAGGCTACAGATGCTCGGCGCTGCTACT 269
DB 472 CCTGCTGAGAAAGTGGCCGTGTTTTCATGACGAGATTGGCTAGTCCCTGCAGTCT 531
OY 270 GCACAGAAAAATCATCTCTGTGTGACCTGAAAGTGGAGACATTTCTGTGTCTCT 329
DB 532 GCATGAGCAAAATCTCTCATCTGTGATGAAACGACAGACATCTCTGTGAGCTCTT 591
OY 330 TGAGGTCAGAGAGACATCAACATCAAGCTATCTGACTAGGGATTCGAGGCAATAT 389
DB 592 -----GGAGAGCCCACTGAAACTGGCAAGCTTTGGCTTGGCCAGACATCTC 642
OY 390 CCATGAGGCGGCTTAGGCGTG---GAGGCGACICTGTGCTACAGGCGCCAGAGATCAG 446
DB 643 CCGCTGAGAGAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702
OY 447 GCTCGCATTTATATAGATGAGAGAGTATGATGATGCTCTCTATGAGAAATGCTCTTACGA 506
DB 703 GTGTGCGGCGGAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 762

OY 507 GTGCTGTGAGAGACAGCCGCC 527
DB 763 AGCCCTTGTGGGAGCCGCC 783

Search completed: April 15, 2003, 07:05:58
Job time : 416 secs

	Best Local Similarity	100.0%; Pred. No. 0;	Matches 728; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
QY	1 MLRLRATDAMKNSEPRROEASMLHALQHPCIVALLIGISIHPLCFATELAPLSLNTVLS	60					
DG	1 MLRLRATDAMKNSEPRROEASMLHALQHPCIVALLIGISIHPLCFATELAPLSLNTVLS	60					
QY	61 ENADSSPIPLGHMLTOKIAVAQIASGLAYLLKKNIIFEDDKSDNILWMSLDVKRHINKL	120					
DG	61 ENADSSSPIPLGHMLTOKIAVAQIASGLAYLLKKNIIFEDDKSDNILWMSLDVKRHINKL	120					
QY	121 SDYGSRQSFEHGAAGVEGTGYOAPETIRPTIYDEKDMRSYGAVLYTELLSGORPALGH	180					
DG	121 SDYGSRQSFEHGAAGVEGTGYOAPETIRPTIYDEKDMRSYGAVLYTELLSGORPALGH	180					
QY	181 HQLOIAKLKSXGIPPVLGQPEPEVOFRRLQALAMECDTKPERKPRLATSVSQMNDPFPAT	240					
DG	181 HQLOIAKLKSXGIPPVLGQPEPEVOFRRLQALAMECDTKPERKPRLATSVSQMNDPFPAT	240					
QY	241 FMYELCGCKOTAFSSSQOEYTVVFMDSKEESRNVTYVTEKGLEMYORMCCPGMKVSCQ	300					
DG	241 FMYELCGCKOTAFSSSQOEYTVVFMDSKEESRNVTYVTEKGLEMYORMCCPGMKVSCQ	300					
QY	301 LQVGRSLMTAEEDQKIYYITLAKGCPLNTPOQALDTPRAVYTCFLAAPPYIKNSVLYLAGL	360					
DG	301 LQVGRSLMTAEEDQKIYYITLAKGCPLNTPOQALDTPRAVYTCFLAAPPYIKNSVLYLAGL	360					
QY	361 ADGLVAVPPVVRGPTRPKDCSYLCSTHTANRSKFSTADEBARONPYPKAMEVYNSGSEVMY	420					
DG	361 ADGLVAVPPVVRGPTRPKDCSYLCSTHTANRSKFSTADEBARONPYPKAMEVYNSGSEVMY	420					
QY	421 SNGBGLVTDCASELICRLEBPYMAPSWMYSVCSSEGRGEVYWCILDCKANSILVMYST	480					
DG	421 SNGBGLVTDCASELICRLEBPYMAPSWMYSVCSSEGRGEVYWCILDCKANSILVMYST	480					
QY	481 TYOLCARFECGVPSPLRPMFPVRPLDTPPPAASHANKVVBGDSDIADVSIYISELTQTQ	540					
DG	481 TYOLCARFECGVPSPLRPMFPVRPLDTPPPAASHANKVVBGDSDIADVSIYISELTQTQ	540					
QY	541 ILIHQESTLTDVCSMSYSSTSPPROARSPSSLSPASSSVPFSTDCEDEDMLTTPGAA	600					
DG	541 ILIHQESTLTDVCSMSYSSTSPPROARSPSSLSPASSSVPFSTDCEDEDMLTTPGAA	600					
QY	601 SDRSEHDLPMDGETFSQHQAANKILAVRDLIIMVFRRGDVIVIGLEKDESEAQRGVYAV	660					
DG	601 SDRSEHDLPMDGETFSQHQAANKILAVRDLIIMVFRRGDVIVIGLEKDESEAQRGVYAV	660					
QY	661 LKARELTPHGVLVDAAVAAKDTVCTFENENTEMCLAWRGWARGAREPFIFFQSYIELGRL	720					
DG	661 LKARELTPHGVLVDAAVAAKDTVCTFENENTEMCLAWRGWARGAREPFIFFQSYIELGRL	720					
QY	721 EACSTRKR 728						
DG	721 EACTRRKR 728						
<hr/>							
RESULT 2							
ID	ABG64871 standard; Protein: 706 AA.						
XX	ABG64871;						
XX	27-AUG-2002 (first entry)						
XX							
XX	Human albumin fusion protein #1546.						
KM	Albumin fusion protein; therapeutic protein X; human albumin; HA;						
KM	human serum albumin; HSA; cancer; reproductive disorder;						
KM	digestive disorder; immune disorder; endocrine disorder;						
KM	hematopoietic disorder; neural disorder; connective disorder;						
KM	cystostatic; antifertility; antiinflammatory; antitumor;						
KM	immunomodulator; anti-HIV; antididiabetic; haemostatic; neotropic;						
KM	neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;						

KW osteopathic; antiarthritic.
 OS Homo sapiens.
 OS Synthetic.
 PN WO200171137-A1.
 PD 18-OCT-2001.
 PD 18-OCT-2001.
 PF 12-APR-2001; 2001WO-US11988.
 PR 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Haseltine WA;
 DR WPI; 2002-010886/01.
 XX
 XX
 PT New fusion protein for treating disease e.g. diabetes comprises an
 PT albumin fused to a therapeutic protein -
 PS Claim 1; Page 1564-1566; 2102pp; English.
 CC The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or
 CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its biological
 CC in vitro/in vivo activity. The protein is useful for treating and
 CC diagnosing disorders such as cancer, reproductive disorders, digestive
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
 CC (e.g. diabetes), hematopoietic disorders, neural disorders
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
 CC (e.g. osteoporosis, arthritis). ABG633726-ABG65518 represent albumin
 CC fusion proteins of the invention.
 SQ Sequence 706 AA;
 Query Match 97.1%; Score 3712; DB 23; Length 706;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 MLEALQHPCTIVALLIGISIHPLCFALIELAPLSLNTVLSNNAROSSFTPLGHMTOKIAYQ 82
 DB 1 MLEALQHPCTIVALLIGISIHPLCFALIELAPLSLNTVLSNNAROSSFTPLGHMTOKIAYQ 60
 QY 83 IASGLATLHKNNITFCOLKSDNLLWSLDVKEHNINIKLSDYGISROSPFHGALGVEGTPG 142
 DB 61 IASGLATLHKNNITFCOLKSDNLLWSLDVKEHNINIKLSDYGISROSPFHGALGVEGTPG 120
 QY 143 YQAPEIRPRIYDEKVMESYGMVLYELLSGORPALGHOLIAKRSKGRIPVLOGPEE 202
 DB 121 YQAPEIRPRIYDEKVMESYGMVLYELLSGORPALGHOLIAKRSKGRIPVLOGPEE 180
 QY 203 VQFRRLQALMECWDTKPEKRPPLALSVSQKDPFATPMYELCCGOTAFESSQGEYTT 262
 DB 181 VQFRRLQALMECWDTKPEKRPPLALSVSQKDPFATPMYELCCGOTAFESSQGEYTT 240
 QY 263 VVEFDGSEESNNTVVTETKGLMEVQRKCGKAKYSCOLQVORSLMATEOKIYITLTK 322
 DB 241 VVEFDGSEESNNTVVTETKGLMEVQRKCGKAKYSCOLQVORSLMATEOKIYITLTK 300
 QY 323 GACGLNTPQOALDTPAVVTCGLAVPIKKNSTYVLAGLADGLVAVFPVRCGPDKSCSYL 382
 DB 301 GACGLNTPQOALDTPAVVTCGLAVPIKKNSTYVLAGLADGLVAVFPVRCGPDKSCSYL 360
 QY 383 CSHTANKSFSTIADEDARONPYPVKAMEVYNSGSEVWYNSGPGLLVLDCALEICRRLPEP 442

DB 361 CSHTANKSFSTIADEDARONPYPVKAMEVYNSGSEVWYNSGPGLLVLDCALEICRRLPEP 420
 QY 443 YMAPSMVTSVYSCSEGRGGEVWCLDDKANSLYVHSTTYQLCARFCGVSPFLRDMFPV 502
 DB 421 YMAPSMVTSVYSCSEGRGGEVWCLDDKANSLYVHSTTYQLCARFCGVSPFLRDMFPV 480
 QY 503 RPLDTEPPAASHANKVPEGDSIADVSTMSSELTQIILHOBESLTDICSMNSYSSSP 562
 DB 481 RPLDTEPPAASHANKVPEGDSIADVSTMSSELTQIILHOBESLTDICSMNSYSSSP 540
 QY 563 RQARSPSSLPSSPASSSSPESTDCEDSDMLHTPGAASDRSHDLTPMDGETFSQHLOA 622
 DB 541 RQARSPSSLPSSPASSSSPESTDCEDSDMLHTPGAASDRSHDLTPMDGETFSQHLOA 600
 QY 623 VKILAVRDLIWPBRGGDVIVIGLEKDSKQGRVIAVLKARELTPHGVLDVAAYAKDT 682
 DB 601 VKILAVRDLIWPBRGGDVIVIGLEKDSKQGRVIAVLKARELTPHGVLDVAAYAKDT 660
 QY 683 VVCTFENENTEMCLAVRWGKAREFDIFQSYELGRLEACTKRRR 728
 DB 661 VVCTFENENTEMCLAVRWGKAREFDIFQSYELGRLEACTKRRR 706
 RESULT 3
 AAE21217
 ID AAE21217 standard; Protein; 706 AA.
 AC AAE21217;
 XX
 DT 01-JUL-2002 (first entry)
 DE Human gene 4 encoded secreted protein HDDMW90, SEQ ID NO:82.
 XX
 KW Human; secreted protein; immune disorder; antiallergic; antirheumatic;
 KW rheumatoid arthritis; breast neoplasia; breast cancer; antiarthritic;
 KW neurological disease; Alzheimer's disease; Parkinson's disease; trauma;
 KW Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania;
 KW antiinflammatory; ophthalmological; dermatological; immunostimulatory;
 KW immunomodulatory; immunosuppressive; antibacterial; angiodystatic;
 KW gene therapy; autoimmune disease; Huntington's disease; meningitis;
 KW demyelinating disease; peripheral neuropathy; congenital malformation;
 KW spinal cord injury; peripheral neuropathy; ischemia; perception;
 KW multiple sclerosis; infection; haemorrhage; schizophrenia; dementia;
 KW depression; panic disorder; learning disability; AIDS; feeding disorder;
 KW hyperproliferative disorder; sleep pattern; cardiovascular disorder;
 KW reproductive disorder; digestive system disorder; behavioural disorder.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT Protein 25..706 /label= Signal_peptide
 FT /label= Mature_secreted_protein
 XX
 PN WO200216390-A1.
 PD 28-FEB-2002.
 PF 17-JAN-2001; 2001WO-US01435.
 PR 18-AUG-2000; 2000US-226282P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
 PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M,
 PI Ni J;
 DR WPI; 2002-304113/34.
 DR N-PDB; AAD337718.
 XX
 PT An isolated nucleic acid molecule (1) comprising a polynucleotide which

PT encodes a polypeptide useful in the diagnosis and treatment of
 PT disorders e.g. immune disorders -
 XX
 PS Claim 11; Page 491-493; 534pp; English.

CC AAD33692-AAD33736 represent cDNAs corresponding to 21 human secreted
 CC protein genes, and AAE21191-AAE21235 represent the proteins they encode.
 CC AAE21136-AAE21280 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 21 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC immune or autoimmune diseases e.g. AIDS (acquired immune deficiency
 CC syndrome), asthma, anaemia and rheumatoid arthritis, breast neoplasia
 CC and breast cancer, neurological diseases e.g. Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, Tourette syndrome,
 CC meningitis, demyelinating disease, peripheral neuropathies, neoplasia,
 CC trauma, congenital malformations, spinal cord injuries, toxic
 CC neuropathies induced by neurotoxins, peripheral neuropathies, multiple
 CC sclerosis, ischaemia and infarction, haemorrhages, schizophrenia, mania,
 CC dementia, depression, panic disorder, learning disabilities, ALS,
 CC altered behaviour e.g. disorders in feeding, sleep patterns, balance
 CC and perception, encephalitis, disorders in cardiovascular, neural/
 CC sensory, reproductive and digestive systems, behavioural disorders and
 CC hyperproliferative disorder. The present sequence represents a human
 CC secreted protein of the invention.

XX Sequence 706 AA;

Query Match 97.1%; Score 3712; DB 23; Length 706;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 MLHAIQHPCIVAIIGISIHPLCFALFALPLSLNTVLSSENRDSSFIFLGHMLTKRIAYQ 82
 DB 1 MLHAIQHPCIVAIIGISIHPLCFALFALPLSLNTVLSSENRDSSFIFLGHMLTKRIAYQ 60

QY 83 IASGLAVYHKKNIIFCDLAKSNILYMSLDVYEHNIKISDYGISQSHHEGALGYEGTGP 142
 DB 61 IASGLAVYHKKNIIFCDLAKSNILYMSLDVYEHNIKISDYGISQSHHEGALGYEGTGP 120

QY 143 YQAEIRPRIYVDEKVDYFSGYVLYELLSGORPALGHNOIYAKKISKGRIPYVGDEE 202
 DB 121 YQAEIRPRIYVDEKVDYFSGYVLYELLSGORPALGHNOIYAKKISKGRIPYVGDEE 180

QY 203 VQFRRLQALMECWDTPREKRPALSYVSKMDPTFATFMTLCCGKQTAFFSSGQERYT 262
 DB 181 VQFRRLQALMECWDTPREKRPALSYVSKMDPTFATFMTLCCGKQTAFFSSGQERYT 240

QY 263 VVFDGKRESENNYVYVNTYKGLMEVORBCGCPGKMYSCOLOVQRSMWTDEOKITTYTLK 322
 DB 241 VVFDGKRESENNYVYVNTYKGLMEVORBCGCPGKMYSCOLOVQRSMWTDEOKITTYTLK 300

QY 323 GMCPPLNTFQOALDTPAVVTCFLAVYIKKNSYLYLAGLADGLVAVFPYVGRTPKDCSYL 382
 DB 301 GMCPPLNTFQOALDTPAVVTCFLAVYIKKNSYLYLAGLADGLVAVFPYVGRTPKDCSYL 360

QY 383 CSHTANRSKFSIADDAQNFYVYKAMEVNVSGSEWTSNGBGLLVLDCALEICRLRLEP 442
 DB 361 CSHTANRSKFSIADDAQNFYVYKAMEVNVSGSEWTSNGBGLLVLDCALEICRLRLEP 420

QY 443 YMASSMTSVYVCSSEGRSEEVVYVCCDDKANSILVYVHSTYQLCARFEGVSPPLDMPPV 502
 DB 421 YMASSMTSVYVCSSEGRSEEVVYVCCDDKANSILVYVHSTYQLCARFEGVSPPLDMPPV 480

QY 503 RPLDTEPPAASHTANPKYPEGDSIADYSIMTSEELGTQILLHQBESLTYVCSMSSSSPP 562
 DB 481 RPLDTEPPAASHTANPKYPEGDSIADYSIMTSEELGTQILLHQBESLTYVCSMSSSSPP 540

QY 563 RQAARSPSLSPSPASSSSVFPSTDCEDSDMLHTPGGAASDRSEHDLTPMDGETFSQHLOA 622

DB 541 RQAARSPSLSPSPASSSSVFPSTDCEDSDMLHTPGGAASDRSEHDLTPMDGETFSQHLOA 600
 QY 623 VKIIAVNRLIVPPRGGVYIYIGLEKSEAGRGVIAVTKARELTPPGVLYDAVVAKDT 682
 DB 601 VKIIAVNRLIVPPRGGVYIYIGLEKSEAGRGVIAVTKARELTPPGVLYDAVVAKDT 660
 QY 683 VVCFEENETEMCAVWNGMGARFEDIFYOSYEELGRLEACTRRRR 728
 DB 661 VVCFEENETEMCAVWNGMGARFEDIFYOSYEELGRLEACTRRRR 706

RESULT 4

ID ABG64872 standard; Protein: 706 AA.

AC ABG64872;

DT 27-AUG-2002 (first entry)

DE Human albumin fusion protein #1547.

XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cytostatic; antifertility; antiinflammatory; antitumor;
 KW immunomodulator; anti-HIV; anti-diabetic; haemostatic; neurotropic;
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antiarthritic.

OS Homo sapiens.

OS Synthetic.

PN WC200177137-AL.

PD 18-OCT-2001.

PF 12-APR-2001; 2001WO-0511988.

PR 12-APR-2000; 2000US-229358P.

PR 25-APR-2000; 2000US-199384P.

PR 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI: 2002-010886/01.

PT New fusion protein for treating disease e.g. diabetes comprises an

PT albumin fused to a therapeutic protein -

PS Claim 1; Page 1566-1569; 2102pp; English.

CC The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or
 CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its biological
 CC in vitro/in vivo activity. The protein is useful for treating and
 CC diagnosing disorders such as cancer, reproductive disorders, digestive
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
 CC (e.g. diabetes), haematopoietic disorders, neural disorders
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
 CC fusion proteins of the invention.

SO Sequence 706 AA;

Query Match 96.9%; Score 3705; DB 23; Length 706;

	Best Local Similarity	99.9%	Pred. No. 0;	Matches 705;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	23	MIHALQPCIVALLGISIHPLCFALFELAPLASSLNTVLSNARDSSFIPLGHMLTKRIAYQ	82					
Db	1	MIHALQPCIVALLGISIHPLCFALFELAPLASSLNTVLSNARDSSFIPLGHMLTKRIAYQ	60					
QY	83	IASSGLAYLHKNNITIFCDLSDNIIILWMSLDYKRNHINIKXLSVDGISIRQSFHEGALGVGCTG	142					
Db	61	IASSGLAYLHKNNITIFCDLSDNIIILWMSLDYKRNHINIKXLSVDGISIRQSFHEGALGVGCTG	120					
QY	143	YQAPEIRPRIYDEFEVDMEFSGMYVELYELISGQRPALIGHQLOIAKKLSIGIRPIYLGQPEE	202					
Db	121	YQAPEIRPRIYDEFEVDMEFSGMYVELYELISGQRPALIGHQLOIAKKLSIGIRPIYLGQPEE	180					
QY	203	VOFRRLQALAMECKWTKPEKRPALASVVSQMKDPFFATPIYTELCCGKQTAFFSSQGEYT	262					
Db	181	VOFRRLQALAMECKWTKPEKRPALASVVSQMKDPFFATPIYTELCCGKQTAFFSSQGEYT	240					
QY	263	VFWMGKEESRMYVYVNNKELMEYORMCCPMKATVSCOLQVRSMTATAEQKITYITLK	322					
Db	241	VFWMGKEESRMYVYVNNKELMEYORMCCPMKATVSCOLQVRSMTATAEQKITYITLK	300					
QY	323	GMCPPLNTPOQALADTPAVVYTCFLAVPIYIKNSLYLAGLADGLVAAPVYVKGTPKDCSTYL	382					
Db	301	GMCPPLNTPOQALADTPAVVYTCFLAVPIYIKNSLYLAGLADGLVAAPVYVKGTPKDCSTYL	360					
QY	383	CSHTANRSKFSIADBDARQNPYPVAKMEVYVNSGSEVWYSGNGPLVYIDCASLEICRLEP	442					
Db	361	CSHTANRSKFSIADBDARQNPYPVAKMEVYVNSGSEVWYSGNGPLVYIDCASLEICRLEP	420					
QY	443	YVABPMVTVSVVCSSEGRGEVYVWCLDDKANKSLVMHSTTYQLCARFCGVPSPRLDMFPY	502					
Db	421	YVABPMVTVSVVCSSEGRGEVYVWCLDDKANKSLVMHSTTYQLCARFCGVPSPRLDMFPY	480					
QY	503	RLDIEPPASHTANPKVVEGDSIADVSLMYSBEELGTOLLHOESLTDVCSMSYSVSSPP	562					
Db	481	RLDIEPPASHTANPKVVEGDSIADVSLMYSBEELGTOLLHOESLTDVCSMSYSVSSPP	540					
QY	563	ROAASPSSLPSSPASSSSVPESTDCESDMLHTPGAADDRSHDLTPMDGEFFSOHLQA	622					
Db	541	ROAASPSSLPSSPASSSSVPESTDCESDMLHTPGAADDRSHDLTPMDGEFFSOHLQA	600					
QY	623	VKIIAVRDLIWPFRGGDIYVIGLEKDSERQGRVIAVLKARLTPHGVLVDAAYAAKDT	682					
Db	601	VKIIAVRDLIWPFRGGDIYVIGLEKDSERQGRVIAVLKARLTPHGVLVDAAYAAKDT	660					
QY	683	VYCTFENENTEMCLAVYRGMGAREPDIYFOSVYELGRLEACTCKRR	728					
Db	661	VYCTFENENTEMCLAVYRGMGAREPDIYFOSVYELGRLEACTCKRR	706					
RESULT 5								
AAE21194	194							
XX	AAE21194	standard; Protein; 706 AA.						
AC	AAE21194;							
XX								
XX	01-JUL-2002	(first entry)						
DE								
XX	Human gene 4	encoded secreted protein HDDMW90, SEQ ID NO:59.						
XX								
KW	Human; secreted protein; immune disorder; antiallergic; antirheumatic;							
KW	rheumatoid arthritis; breast neoplasia; breast cancer; antiarthritic;							
KW	neurological disease; Alzheimer's disease; Parkinson's disease; trauma;							
KW	Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania;							
KW	antiinflammatory; ophthalmological; dermatological; immunostimulatory;							
KW	immunomodulatory; immunosuppressive; antibacterial; antiposrotic;							
KW	gene therapy; autoimmune disease; Huntington's disease; meningitis;							
KW	demyelinating disease; peripheral neuropathy; congenital malformation;							
KW	spinal cord injury; peripheral neuropathy; ischemia; perception;							
KW	multiple sclerosis; infarction; haemorrhage; schizophrenia; dementia;							
KW	depression; panic disorder; learning disability; AIDS; feeding disorder;							

XX	Homo sapiens.
OS	
XX	
FM	Key
FT	Peptide
FT	Protein
FT	/note= "Human mature secreted protein"
FN	
XX	MO200216390-A1.
PD	
XX	28-FEB-2002.
PF	
XX	17-JAN-2001; 2001WO-US01435.
XX	
PR	18-AUG-2000; 2000US-226282P.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
P1	Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS,
P1	Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M,
P1	Ni J;
DR	WPI: 2002-304113/34.
DR	N-PSDB; AAD33695.
PT	An isolated nucleic acid molecule (I) comprising a polynucleotide which
PT	encodes a polypeptide useful in the diagnosis and treatment of
PT	disorders e.g. immune disorders -
PS	Claim 11; Page 472-474; 534pp; English.
XX	
CC	AAD33692-AAD33736 represent cDNAs corresponding to 21 human secreted
CC	protein genes, and AAE21191-AAE21235 represent the proteins they encode.
CC	AAE21236-AAE21280 represent human secreted protein fragments. The genes
CC	and their corresponding secreted proteins are useful for preventing,
CC	treating or ameliorating medical conditions, e.g., by protein or gene
CC	therapy. Pathological conditions can be diagnosed by determining the
CC	amount of the new protein in a sample or by determining the presence of
CC	mutations in the new genes. Specific uses are described for each of the
CC	21 genes, based on the tissues in which they are most highly expressed,
CC	and include developing products for the diagnosis or treatment of
CC	immune or autoimmune diseases e.g. AIDS (acquired immune deficiency
CC	syndrome), asthma, anaemia and rheumatoid arthritis, breast neoplasia
CC	and breast cancer, neurological diseases e.g. Alzheimer's disease,
CC	Parkinson's disease, Huntington's disease, Tourette syndrome,
CC	meningitis, demyelinating disease, peripheral neuropathies, neoplasia,
CC	trauma, congenital malformations, spinal cord injuries, toxic
CC	neuropathies induced by neurotoxins, peripheral neuropathies, multiple
CC	sclerosis, ischemia and infarction, haemorrhages, schizophrenia, mania,
CC	dementia, depression, panic disorder, learning disabilities, ALS,
CC	altered behaviours e.g. disorders in feeding, sleep patterns, balance
CC	and perception, encephalitis, disorders in cardiovascular, neural/
CC	sensory, reproductive and digestive systems, behavioural disorders and
CC	hyperproliferative disorder. The present sequence represents a human
CC	secreted protein of the invention.
XX	
SQ	Sequence 706 AA:
Query Match	96.9%; Score 3705; DB 23; Length 706;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 705; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	23 MHAIHQPCIVALLGISIHPLCPALFELAPLASSINLYLSSENARSSFPICGHMLTKRIAYO 82
Db	1 MHAIHQPCIVALLGISIHPLCPALFELAPLASSINLYLSSENARSSFPICGHMLTKRIAYO 60
OY	83 IASGLAYLEKRNKIIFCDLSDNLIWVSLDYVEKEHNIRIKLSDYGISRSQSFEGALGVGTG 142
Db	61 IASGLAYLEKRNKIIFCDLSDNLIWVSLDYVEKEHNIRIKLSDYGISRSQSFEGALGVGTG 120

QY 143 YQAPRIIRIVDEKVMESYGVLYELLSGORPALGHOLQIAKIKLSKIRPVLAGQPEE 202
 DB 121 YQAPRIIRIVDEKVMESYGVLYELLSGORPALGHOLQIAKIKLSKIRPVLAGQPEE 180
 QY 203 VQRRIRQALMEGWDRKPEKRPALSVSQMKDPTFTFTYELCCGQTAFSSQGOEYT 262
 DB 181 VQRRIRQALMEGWDRKPEKRPALSVSQMKDPTFTFTYELCCGQTAFSSQGOEYT 240
 QY 263 VFWMDKEESRNTVYVTEKGLMEVQRMCCPGKVSQCLQVORSIMATDOKIYITLTK 322
 DB 241 VFWMDKEESRNTVYVTEKGLMEVQRMCCPGKVSQCLQVORSIMATDOKIYITLTK 300
 QY 323 GNCPLMTPOQALDTPAVVTCFLAVPYIKKNSYLVLAGLADGLVAFFVNGTFRKDSQSYL 382
 DB 301 GNCPLMTPOQALDTPAVVTCFLAVPYIKKNSYLVLAGLADGLVAFFVNGTFRKDSQSYL 360
 QY 383 CSHTARSKFSIADEDARQNPYPVKAMEVNSGSEWYSNGPGLVYDCALEICRLEP 442
 DB 361 CSHTARSKFSIADEDARQNPYPVKAMEVNSGSEWYSNGPGLVYDCALEICRLEP 420
 QY 443 YMAPSVTVSVCSSEGRGEVWVCLDKANSIYVHSTYQLCARFPGVPSPLRDMFPV 502
 DB 421 YMAPSVTVSVCSSEGRGEVWVCLDKANSIYVHSTYQLCARFPGVPSPLRDMFPV 480
 QY 503 RPLDTEPPASHTANKRPVPGDSIADVSIWYSEELGTQILIHQESLTDYCSMSYSSSP 562
 DB 481 RPLDTEPPASHTANKRPVPGDSIADVSIWYSEELGTQILIHQESLTDYCSMSYSSSP 540
 QY 563 ROAARSPSLPSSPASSSVFSTDCEDSMLTPGAASRSHDLTPMGGEFFSOHLQ 622
 DB 541 ROAARSPSLPSSPASSSVFSTDCEDSMLTPGAASRSHDLTPMGGEFFSOHLQ 600
 QY 623 VKLIARVDLIWRRGGDVIVIGLEKDSQAQRVIAVLAKEBELTPHGVYDAVAAYAKDT 682
 DB 601 VKLIARVDLIWRRGGDVIVIGLEKDSQAQRVIAVLAKEBELTPHGVYDAVAAYAKDT 660
 QY 683 VYCFEENENTEMCLAVRWGAGAREPDIFFIOSYELGRLAEACTRRR 728
 DB 661 VYCFEENENTEMCLAVRWGAGAREPDIFFIOSYELGRLAEACTRRR 706
 RESULT 6
 AAG67395
 ID AAG67395 standard; Protein: 2014 AA.
 AC AAG67395;
 XX
 DT 13-NOV-2001: (first entry)
 XX
 DE Amino acid sequence of human protein kinase SGK258.
 XX
 KW Human; protein kinase; cancer; immune disease; cardiovascular disease;
 KW brain disease; neuronal disease; Alzheimer's disease; chromosome 15;
 KW Parkinson's disease; multiple sclerosis; metabolic disorder;
 KW peripheral nervous system disease; amyotrophic lateral sclerosis;
 KW infection; ocular disease; migraine; pain; sexual dysfunction;
 KW mood disorder; attention disorder; cognition disorder; hypotension;
 KW hypertension; psychotic disorder; dyskinesia; transplant rejection.
 XX
 OS Homo sapiens.
 PN WO200166594-A2.
 PD 13-SEP-2001.
 XX
 PF 02-MAR-2001: 2001WO-US06838.
 XX
 PR 06-MAR-2000: 2000US-0187150.
 PR 29-MAR-2000: 2000US-0193404.
 PR 13-NOV-2000: 2000US-0247013.
 XX
 PA (SUGEN-) SUGEN INC.
 XX

PI Plowman GD, Whyte D, Manning G, Sudarshanam S, Martinez R;
 XX WPI: 2001-536777/59.
 DR N-PSDB; AAH77994.
 XX
 PT Nucleic acids capable of encoding human polypeptides having a kinase or
 PT kinase-like activity, useful for diagnosing a disease selected from
 PT cancer, cardiovascular disease and neuronal-associated diseases (e.g.
 PT Alzheimer's disease) -
 XX
 PS Claim 7; Fig 2B-C; 201pp; English.
 XX
 CC The present sequence represents a human protein kinase. The gene is
 CC located on chromosome 15. The kinase polypeptides are useful for
 CC diagnosing a disease or disorder selected from cancers (e.g. cancers
 CC of tissues and cancers of hematopoietic origin), immune-related diseases
 CC and disorders, cardiovascular disease, brain or neuronal-associated
 CC diseases (e.g. Alzheimer's disease, Parkinson's disease, multiple
 CC sclerosis), metabolic disorders, peripheral nervous system diseases,
 CC amyotrophic lateral sclerosis, viral infections, infections caused by
 CC prions, infections caused by bacteria, infections caused by fungi,
 CC ocular diseases, migraines, pain, sexual dysfunction, mood disorders,
 CC attention disorders, cognition disorders, hypotension, hypertension,
 CC psychotic disorders, dyskinesias, and organ transplant rejection.
 CC Kinase inhibitors are useful for treating diseases and disorders
 CC described above.
 XX
 SQ Sequence 2014 AA;
 Query Match 93.7%; Score 3579.5; DB 22; Length 2014;
 Best Local Similarity 94.5%; Pred. No. 1,4e-312;
 Matches 688; Conservative 0; Mismatches 1; Indels 39; Gaps 1;
 QY 1 MRLHATATAMKRFSEFRQASMLHALQPCYALIGISIRHCPLELAPSSLNTVVS 60
 DB 1326 MRLHATATAMKRFSEFRQASMLHALQPCYALIGISIRHCPLELAPSSLNTVVS 1385
 QY 61 ENARDSFPIGLHMLTQKIAVQASGLATLHKNNIFCDLKSNNILWMSLDVKEHINIKL 120
 DB 1386 ENARDSFPIGLHMLTQKIAVQASGLATLHKNNIFCDLKSNNILWMSLDVKEHINIKL 1445
 QY 121 SDYGISRQSFHEGALGVECTPGYQAEIRPRIYVDEKVMESYGVLYELLSGORPALGH 180
 DB 1446 SDYGISRQSFHEGALGVECTPGYQAEIRPRIYVDEKVMESYGVLYELLSGORPALGH 1505
 QY 181 HQQIAKIKLSKIRPVLAGQPEEYQFRRLQALMEGWDRKPEKRPALSVSQMKDPTFT 240
 DB 1506 HQQIAKIKLSKIRPVLAGQPEEYQFRRLQALMEGWDRKPEKRPALSVSQMKDPTFT 1565
 QY 241 FMYELCCGKOTAFSSQGOEYTVFWDGKEESRNTVYVTEKGLMEVQRMCCPGKVSQCL 300
 DB 1566 FMYELCCGKOTAFSSQGOEYTVFWDGKEESRNTVYVTEKGLMEVQRMCCPGKVSQCL 1625
 QY 301 LOVQSRSLMATEOKIYITLTKGMCPNLTPQALDTPAVVTCFLAVPYIKKNSYLVLAGL 360
 DB 1626 LOVQSRSLMATEOKIYITLTKGMCPNLTPQALDTPAVVTCFLAVPYIKKNSYLVLAGL 1646
 QY 361 ADGLVAVFPVYVNGTFRKDSQSYLCSHTANRKSIAEDARQNPYPVKAMEVNSGSEWY 420
 DB 1647 ADGLVAVFPVYVNGTFRKDSQSYLCSHTANRKSIAEDARQNPYPVKAMEVNSGSEWY 1706
 QY 421 SNGPGLVYDCALEICRLEPYPMAVMTSVVCSSEGRGEVWVCLDKANSIYVHST 480
 DB 1707 SNGPGLVYDCALEICRLEPYPMAVMTSVVCSSEGRGEVWVCLDKANSIYVHST 1766
 QY 481 TYQLCARFPGVPSPLRDMFPVRLDTEPPASHTANKRPVPGDSIADVSIWYSEELGTQ 540
 DB 1767 TYQLCARFPGVPSPLRDMFPVRLDTEPPASHTANKRPVPGDSIADVSIWYSEELGTQ 1826
 QY 541 ILIHQESLTDYCSMSYSSSPRQAAARSPSLPSSPASSSVFSTDCEDSMLHTPGAA 600
 DB 1827 ILIHQESLTDYCSMSYSSSPRQAAARSPSLPSSPASSSVFSTDCEDSMLHTPGAA 1886

QY	601	SDSSEHDLPMDSEFSSOHAKVLAAYVDLLMPVRGGDVIIVLEKDSKQNRVAV	660
DB	1887	SDSSEHDLPMDSEFSSOHAKVLAAYVDLLMPVRGGDVIIVLEKDSKQNRVAV	1946
QY	661	LKARELPHGVLDAAVVAADYVVCFFENENEMCLAVVRGNGAREPDIYQSYELGRL	720
DB	1947	LKARELPHGVLDAAVVAADYVVCFFENENEMCLAVVRGNGAREPDIYQSYELGRL	2006
QY	721	EACTRRR 728	
DB	2007	EACTRRR 2014	
RESULT 7			
AAE24137			
ID	AAE24137	standard; Protein; 2014 AA.	
AC	AAE24137;		
XX			
DT	23-SEP-2002	(first entry)	
DE			
XX		Human kinase (PKIN)-8 protein.	
XX			
KW	Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;		
KW	acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;		
KW	asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;		
KW	development; hepatitis; cardiovascular; hypertension; drug screening;		
KW	myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;		
KW	fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;		
KW	hypercholesterolemia; obesity; gene therapy; cytostatic; anti-hiv;		
KW	neuroprotective; hepatotropic; hypotensive; cardiac; nephrotropic;		
KW	hyperlipidaemia; enzyme.		
OS	Homo sapiens.		
XX			
FT			
FT	Key	Location/Qualifiers	
FT	Domain	1280..1549	
FT		/note="Eukaryotic protein kinase domain"	
FT	Domain	1295..1549	
FT		/note="Protein kinase domain"	
FT	Domain	1367..1387	
FT		/note="transmembrane domain"	
PN	W0200233099-A2.		
XX			
PD	25-APR-2002.		
XX			
PF	20-OCT-2001;	2001MO-US47728.	
XX			
PR	20-OCT-2000;	2000US-242410P.	
PR	27-OCT-2000;	2000US-244068P.	
PR	03-NOV-2000;	2000US-245708P.	
PR	09-NOV-2000;	2000US-247672P.	
PR	16-NOV-2000;	2000US-249565P.	
PR	22-NOV-2000;	2000US-252730P.	
PR	01-DEC-2000;	2000US-250807P.	
XX			
PA	(INCYTE GENOMICS INC.		
XX			
PI	Gururajan R, Baughn MR, Nallia NK, Elliott VS, Xu Y, Arvizu C;		
PI	Yao MG, Ramkumar J, Ding L, Tang YT, Hatalla AA, Nguyen DB;		
PI	Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CN;		
PI	Lal PG, Recipon SA, Lu DM, Borowsky ML, Thornton M, Swarnaker A		
PI	Thangavelu K, Khan FA, Ison CH;		
XX			
DR	WPI: 2002-454603/48.		
DR	N-PSDB; AAD38R51.		
XX			
PT	New human kinase polypeptide, for diagnosing, preventing and treating		
PT	cancer, immune system disorders, growth and development disorders,		
PT	cardiovascular disorders and lipid disorders		
XX			
PS	Claim 1; Page 147-152; 210pp; English.		

XX	The invention relates human kinases (PKIN) and their corresponding
CC	nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
CC	treating and preventing cancer, an immune system disorder (e.g.,
CC	acquired immune deficiency syndrome (AIDS), Addison's disease, allergy,
CC	asthma, atherosclerosis, multiple sclerosis, psoriasis), disorders
CC	affecting growth and development (e.g., arteriosclerosis, cirrhosis,
CC	hepatitis), cardiovascular disorder (e.g., hypertension, myocardial
CC	infarction, Goodpasture's syndrome), and a lipid disorder (e.g., fatty
CC	liver, Gaucher's disease, Nieman-Pick's disease, hypercholesterolemia,
CC	hyperlipidaemia, obesity), and for assessing the effects of exogenous
CC	compounds. Anti-PKIN antibody is useful in a diagnostic test for a
CC	condition or a disease associated with the expression of PKIN in a
CC	biological sample. A composition comprising PKIN or an agonist or
CC	antagonist of PKIN is useful for treating a disease or condition
CC	associated with decreased or increased expression of functional PKIN.
CC	PKIN is useful in a number of drug screening techniques and to analyse
CC	the proteome of a tissue or cell type. PKIN DNA is useful for creating
CC	knockin humanised animals or transgenic animals to model human diseases,
CC	and in somatic or germline gene therapy. The present sequence is human
CC	PKIN protein.
CC	
XX	
SQ	Sequence 2014 AA;
	Query Match 93.7%; Score 3579.5; DB 23; Length 2014;
	Best Local Similarity 94.5%; Pred. No. 14e-312;
	Matches 688; Conservative 0; Mismatches 1; Indels 39; Gaps 1.
OY	1 MLRHRAIDAMKNSEFPEQASMLHALQHPICVALIGISHPRLCFALDELAPLSLNTVLS 60
DB	1326 MLRHRAIDAMKNSEFPEQASMLHALQHPICVALIGISHPLCFALDELAPLSLNTVLS 1385
OY	61 ENARDSSTPIGHLTOKIAIQAIOIASGLAYLHKNNIIIFCDLKSDNLIWMSLDVKHENITL 120
DB	1386 ENARDSSTPIGHMTOKIAIQAIOIASGLAYLHKNNIIIFCDLKSDNLIWMSLDVKHEHINTL 1445
OY	121 SDYGISRGSFHFGALGVGTSGYOAPETRPRIYVDYEKVMSFYGNVLVELLSGORPALGH 180
DB	1446 SDYGISRGSFHFGALGVGTSGYOAPETRPRIYVDYEKVMSFYGNVLVELLSGORPALGH 1505
OY	181 HOLQIAKLTKSKIRPVLGQPEEVOFRRLQALAMECNDTRPEKRPLALSVMQMDTFPAT 240
DB	1506 HOLQIAKLTKSKIRPVLGQPEEVOFRRLQALAMECNDTRPEKRPLALSVMQMDTPTAT 1565
OY	241 FMYELCCGKQTAFSSSOGGYTVFWMDGKEESRNITVTTEKGLMEVQRCCPGMKVSQC 300
DB	1566 FMYELCCGKQTAFSSSOGGYTVFWMDGKEESRNITVTTEKGLMEVQRCCPGMKVSQC 1625
OY	301 LQVQSLMTPATDQKIITYTLTKMGCPLPNPQDALDPFAVVTCFLANPVYIKNSLYLAGI 360
DB	1626 LQVQSLMTPATE-----N-SLYLAGI 1646
OY	361 ADGLAVAFPVVGTEPRKDSCTSYLCSHTANKSFESIADEDARONPYPYKAMEVNVSSGEVNY 420
DB	1647 ADGLVAFPVVGTEPRKDSCTSYLCSHTANKSRKSINADEDARONPYPYKAMEVNVSSGEVNY 1706
OY	421 SNGPGLLVIDCASLEICRRLEPYMAPSWTSTVYCSESGEGEEYWCLODKANSIVYHST 480
DB	1707 SNGPGLLVIDCASLEICRRLEPYMAPSWTSTVYCSESGEGEEYWCLODKANSIVYHST 1766
OY	481 TYQLCARFCYGVPSPRIDEPVRPLDTDEPPASHTNPNVPBEDSIADVSIWSEELGRQ 540
DB	1767 TYQLCARFCYGVPSPRIDEPVRPLDTDEPPASHTNPNVPBEDSIADVSIWSEELGRQ 1826
OY	541 ITHIESSLIDVYSMSMSYSSSPROAKRSPSLPSPPASSSVPFSTDCESDMLHTPGNA 600
DB	1827 ITHIESSLIDVYSMSMSYSSSPROAKRSPTSLSPPASSSVPFSTDCESDMLHTPGAA 1886
OY	601 SDRSEHDITPMDCGETFSOHLQAVKITAVRDLLWPFRGGDVYIGLEKSEAORGVYIAV 660
DB	1887 SDRSEHDITPMDCGETFSOHLQAVKITAVYNDLTWVPRGGGVYIIGLEKSGAORGVYIAV 1946
OY	661 LKARELIPHGVLDAVAVKOTVCTCFENMENTEMCLAVMKRCGRANDEFIFYCYTEHLGRI 720

|||||
Db 1947 LKARLTGHTGVLVDAVAAYAKDTVCTFENENETMCIAVWGGAREFDIFYOSYELGRL 2006
Oy 721 EACTRRR 728
|||||
Db 2007 EACTRRR 2014

RESULT 8

ABG08051
ID ABG08051 standard; Protein: 809 AA.

AC ABG08051;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8042.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; NAs72238.

PT New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.

PS Claim 20; SEQ ID NO 38410; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABG00010-ABG30377 represent novel human
diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pcl_sequences.

SC Sequence 809 AA;

Query Match 78.4%; Score 2997; DB 22; Length 809;
Best Local Similarity 93.3%; Pred. No. 1,1e-260;
Matches 584; Conservative 4; Mismatches 26; Indels 12; Gaps 4;

Oy 1 MNRHRLATDANKNESEFROEASMLHALQPCIALIGISIHPLCFALAPLSSINTVLS 60
Db 149 MNRHRLATDANKNESEFROEASMLHALQPCIALIGISIHPLCFALAPLSSINTVLS 208
Oy 61 ENARDSFTPLGHMLTQKIAVOIASGLAYLHKNNIIFCDLSKSDNLVSLDKHEINIKL 120
Db 209 ENARDSFTPLGHMLTQKIAVOIASGLAYLHKNNIIFCDLSKSDNLVSLDKHEINIKL 268
Oy 121 SDYGISRSFHEGALGVGTGQAPETRPRIYVDEKDMNSYGVLYELLSSGRRPALGH 180
Db 269 SDYGISRSFHEGALGVGTGQAPETRPRIYVDEKDMNSYGVLYELLSSGRRPALGH 328
Oy 181 HOLOIAKLSKIRPVLCQPEEYQPRRLQALMECCMDTKPERKRLATSVSOMKDPTE-- 238
Db 329 HOLOIAKLSKIRPVLCQPEEYQPRRLQALMECCMDTKPERKRLATSVSOMKAPDPC 388
Oy 239 --ATFMYELCCGQTAFFSSGQETVVFND-GKEESRNTYVNTENGMEVORMCCPG 294
Db 369 LHVXTVME---ADGLLIPGCVHGVIGWNRGVQELHG--GNTKGLMEVORMCCPG 442
Oy 295 MKVSCQLQVORSIMWTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLAVPYIKNSY 354
Db 443 MKVSCQLQVORSIMWTATEDOKIYITLKGMCPLNTPQALDTPAVVTCFLAVPYIKNSY 502
Oy 355 LVLAGIADGLVAVFPVVRGTGTPKSCSYLCSHTANRSKFSIADEDARQNPYVKAEEVNS 414
Db 503 LVLAGIADGLVAVFPVVRGTGTPKSCSYLCSHTANRSKFSIADEDARQNPYVKAEEVNS 562
Oy 415 GSEWVYNSGPGILVTDCALEICRRLPEYMAPSVYTSVYVCSGEGEEVYVWCLDDKANSL 474
Db 563 GSEWVYNSGPGILVTDCALEICRRLPEYMAPSVYTSVYVCSGEGEEVYVWCLDDKANSL 622
Oy 475 VMYHSTTYQOLCARVFCGVPSPLRDMFPVPRDTEPPASHTANPVVPGDSIADVSIMYS 534
Db 623 VMYHSTTYQOLCARVFCGVPSPLRDMFPVPRDTEPPASHTANPVVPGDSIADVSIMYS 682
Oy 535 EELGTQILIHQESLTDYCSMSYSSSPROAARSPSSLPSSSPASSSVYFSTDCEDSDML 594
Db 683 EELGTQILIHQESLTDYCSMSYSSSPROAARSPSSLPSSSPASSSVYFSTDCEDSDML 742
Oy 595 HTPGAASDRSEHDLTPMDGETFSQHL 620
Db 743 HTPGAASDRSEHDLTPMDGETFSQHL 768

RESULT 9

ABB61564
ID ABB61564 standard; Protein: 2308 AA.

AC ABB61564;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 11484.

KW Drosophila: developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.
 DR N-PSDB: ABL05667.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO.11484; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB120511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB157737-AB172072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 CC
 XX
 SQ Sequence 2308 AA;
 Query Match 9.5%; Score 364.5; DB 22; Length 2308;
 Best Local Similarity 23.7%; Pred. No. 1.1e-22;
 Matches 159; Conservative 99; Mismatches 243; Indels 169; Gaps 27;
 QY 2 LRLHATDAMNFESEFROEASMLHALPHCIYALIGISIHPLCFALPLSLNTVLSE 61
 DB 1684 LQH-----SCAYCATARELAVLTLKHPNTVPLVGTICKLALVELAPLGGDALLRH 1738
 QY 62 NARDSFFPLHMLTKRAYOIASGLAYLAKNNITFCCLKSDNIIWGL-----DYKE 114
 DB 1739 YRSGAH--MGPHTFQTLVLAARAIEYLRRIIRDLKSENVLMELPQHTEDSPRN 1796
 QY 115 HINIKLSYGISROSEFGALGVEGTPGYOAPETIRIVYDEKVMFYGMVLYELLSGO 174
 DB 1797 LVHIIKAIYGISRQTPASGAGFGGTGFAPEL---IRYN----- 1834
 QY 175 RPAIGHOQIAKILSKIGRPVLAGPEEYQFR--LQALMECWDTPKPRRLATSVSOM 233
 DB 1835 ---GHE--SIKECILGSRPALNQ--RETQPTCCLDLWVCMHQPRRRPTASQIVSIL 1887
 QY 234 KDPFFATM-----YELCCGK-QTAFSSQGEIVVWDEKESKNTVVTENG- 283
 DB 1888 SAPECIHLDVYAMPSEKIVCGVFOSLVGMGDERCGLIELMPSFGSRIDILDCSPSGS 1947
 QY 284 LMEVORMCC-----PGKKVSCQLQVQNSIMTATEDKIX 317
 DB 1948 LLOCNSICSPQOVAPPKTPENGANSARSAORLPKNNMLCCCLVGERATMGDV9GMLH 2007
 QY 318 IYTLKGMCLPTPOALDTPAVVTCFLAVPIKNSYLVLAGLADGLAVPVVAVGTGPKD 377
 DB 2008 AYSTSTVAHLS--YMLD-PIKSAVISLVYMERIA--RVAAGTNG--RVLVYDAQTQPS 2061
 QY 378 SCST-----LCSHTANSKFSIADED-----ARQNDYP----- 405
 DB 2062 NCARAGSFVLEICSGVLAHACSVVVDGIVELWCGEIAKINIVPLNENGVSCHQALC 2121
 QY 406 -----VKAMEVNSGSEWVYNSNGPLL-----IDCAILEIC-RRLLE 441
 DB 2122 HSEBNLIEYKVAHMCNSHVSCLYPGCMVYQMDVISRIENKILDCSKILPCESELQ 2181
 QY 442 PYMAPSVMTSVVCSSEGEVEVWCLDKANSIVWYHSTYQOLCARYCGVPSPLRDKFP 501
 DB 2182 STADEHNILKC-----QISALAHNSELV-IGTWOCILVAELHTLRP 2225
 QY 502 V---RPLDTEPPASHTANPVVPSGDSIADYSIMSEELGQOILHOSLFDYCSHSSYS 558
 DB 2226 ISVFPEYNEIKSITTLTKDNVP---LIATIGRRY-RSLISRYVDSAESST---KSSAV 2277
 QY 559 SSPPQAAARS 568

DB 2278 STPTHGAAKS 2287
 ||| |||
 RESULT 10
 AAEL6259
 ID AAEL6259 standard; Protein: 656 AA.
 XX
 AC AAEL6259;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human kinase PKIN-5 protein.
 XX
 KW Human; kinase; PKIN-5; cancer; leukemia; adenocarcinoma; osteoarthritis;
 KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
 KW Acquired immune deficiency syndrome; AIDS; Addison's disease; anaemia;
 KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
 KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoporosis;
 KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
 KW rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
 KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
 KW cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
 KW congestive heart failure; ischaemic heart disease; lung tumour; gout;
 KW fatty liver; Niemann-Pick's disease; gene therapy.
 KW
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 13..252
 FT /label= Protein_kinase_domain
 FT Domain 14..197
 FT /label= Protein_kinase_domain
 FT Domain 14..257
 FT /note= "Eukaryotic protein kinase domain"
 FT Domain 14..252
 FT /label= Protein_kinase_domain
 FT Domain 14..253
 FT /label= Protein_kinase_domain
 XX
 WO200196547-A2.
 PD 20-DEC-2001.
 XX
 PF 14-JUN-2001; 2001WO-US19444.
 XX
 PR 15-JUN-2000; 2000US-212073P.
 PR 23-JUN-2000; 2000US-213467P.
 PR 30-JUN-2000; 2000US-215651P.
 PR 07-JUL-2000; 2000US-216605P.
 PR 13-JUL-2000; 2000US-218372P.
 PR 25-AUG-2000; 2000US-228056P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
 PI Gandhi AR, Tribouley CM, Walla NK, Yao MG, Lu DAM, Greenwald SR;
 PI Rankumar J, Griffin JA, Kearney N, Burford N, Nguyen DB, Tang YF;
 PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
 PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
 PI Grether M, Elliott VS, Thangavelu K, Bacteria S, Ison CH;
 XX
 DR WPI: 2002-090207/12.
 DR N-PSDB: AAD26452.
 XX
 PT New polypeptides, useful for diagnosing, treating or preventing
 PT disorders of growth and development, cardiovascular and lipid, and
 PT diseases such as cancer, comprise human kinase polypeptides -
 XX
 PS Claim 1; Page 143-144; 197pp; English.
 CC The invention relates to human kinase PKIN proteins and their
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for

treating a disease or condition associated with decreased expression of PKIN and a composition comprising PKIN antagonist is useful for treating a disease or condition associated with overexpression of PKIN. The disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, Hodgkin's disease), immune disorder (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease, attherosclerosis, anemia, allergies, adult respiratory distress syndrome, autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease, rheumatoid arthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome, bacterial, parasitic, fungal, viral, protozoal and helminthic infections), growth and development disorders (arteriosclerosis, cataracts), cardiovascular disease (arteriovenous fistula, hypertension, vasculitis, aneurysms, congestive heart failure, angina pectoris, myocarditis, ischaemic heart disease, chronic-bronchitis, lung tumours), lipid disorder (fatty liver, Fabry's disease, Niemann-Pick's disease, hypercholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity of a test compound and in gene therapy. The present sequence is human PKIN-5 protein.

Sequence 656 AA;

Query Match 8.6%; Score 327.5; DB 23; Length 656;
Best Local Similarity 26.1%; Pred. No. 3.1e-20;
Matches 101; Conservative 75; Mismatches 168; Indels 43; Gaps 13;

OY 10 AMKNESE-----FROEASMLHALQPCVALIGISIHPLCALAPLSSINTVLSENA 63
DB 33 AKITNKMTSLRLNQLVAVCHLHHPISLSLAAGIRPMLVMEASKSGSLDLRLQODK 92
OY 64 RDSSEPIPLGHMLTKRIAVQIASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKLSY 123
DB 93 AS-----LTRLQNHIALHVAADGLAKXLSAMITTYIDKPHNVLLTLPNNAITAKIDY 147
OY 124 GISROSFEGALGVEGTGYOAPET-RRPIYDEKVDMSGYVLYELL-SGQRPALG-- 179
DB 148 GIAOYCCRMGKITSEGTGFRAPEVARGNVITNOADVYSFGLLYDILTTGGRIVEGLK 207
OY 180 -----HHOIAKIKLSKGRVPLVGPEEVOFRLOALMECDTRKPKRLALSVYSOKMD 235
DB 208 FPNNEDELEIOGKLPDPVKEGCAPE-----WPNVEKLIKOCLKENQERTSAQVFDILNS 263
OY 236 PTFATFWELOCCGKOT---AFFSSOGQETVYVFDG---KEESRNTVAVTEKGLMEV- 287
DB 264 AELVCLTRILLPKNVIVECVATHHNSRNASIMWGCCHTRGQLSFLDLNTEGTSSEV 323
OY 288 -ORMCPCGMKVSQLOVORSIM--TATEDOKIYITTLKGMCPMLTPPOALDTPAVVTCF 343
DB 324 ADSRLICLAL---VHLPEKESWIVSGTOSGFLVINTEDGKRHTLEKMTDS---VTCL 377
OY 344 LAVPYIK--KNSYLVLAGLADGLVAVF 368
DB 378 YCNSEFSKSKOKNLFVLTGADGKLATF 404

RESULT 11
AAU03554
ID AAU03554 standard; Protein: 909 AA.

AC AAU03554;

DT 12-SEP-2001 (first entry)

DE Human protein kinase #54.

KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;

KW metabolic disorder; immune related disease; neurological disorder;

KW neurodegenerative disorder; inflammatory disorder; infectious disease;

OS Homo sapiens.

XX WO200138503-A2.
XX 31-MAY-2001.
XX 22-NOV-2000; 2000WO-US32085.
XX 24-NOV-1999; 99US-0167482.
XX (SUGEN) SUGEN INC.
XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX Flanagan P, Clary D;
XX WPI: 2001-343950/36.
XX N-PDB: AAS06754.

XX Nucleic acids encoding human kinase polypeptides, useful for preventing
XX diagnosing and/or treating e.g. cancer, immune, cardiovascular and
XX neuronal-associated diseases, and microbial infections -
XX Claim 7; Figure 2; 433pp; English.

XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The
XX novel protein kinases have been identified as members of the tyrosine
XX or serine/threonine kinase (PTK and STK) families. The polynucleotides
XX encoding protein kinases and the polypeptides may be used in the
XX prevention, diagnosis and treatment of diseases associated with
XX inappropriate kinase expression. For example, they may be used to treat
XX cancers (especially cancers of haematopoietic origin), cardiovascular
XX disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
XX immune related diseases (e.g. rheumatoid arthritis), neurological
XX disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
XX Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
XX disease (e.g. HIV) and reproductive disorders (e.g. infertility).
XX Additionally, polynucleotides encoding protein kinases may be
XX used for gene therapy and as DNA probes in diagnostic assays.
XX The protein kinase polypeptides may be used as antigens in the production
XX of antibodies against the protein kinases and in assays to identify
XX modulators of protein kinase expression and activity.

Sequence 909 AA;

Query Match 8.6%; Score 327.5; DB 22; Length 909;
Best Local Similarity 26.1%; Pred. No. 5.3e-20;
Matches 101; Conservative 75; Mismatches 168; Indels 43; Gaps 13;

OY 10 AMKNESE-----FROEASMLHALQPCVALIGISIHPLCALAPLSSINTVLSENA 63
DB 419 AVKINKMTSLRLNQLVAVCHLHHPISLSLAAGIRPMLVMEASKSGSLDLRLQODK 478
OY 64 RDSSEPIPLGHMLTKRIAVQIASGLAYLHKNNIIFCDLKSNDILVMSLDVKEHINIKLSY 123
DB 479 AS-----LTRLQNHIALHVAADGLAKXLSAMITTYIDKPHNVLLTLPNNAITAKIDY 533
OY 124 GISROSFEGALGVEGTGYOAPET-RRPIYDEKVDMSGYVLYELL-SGQRPALG-- 179
DB 534 GIAOYCCRMGKITSEGTGFRAPEVARGNVITNOADVYSFGLLYDILTTGGRIVEGLK 593
OY 180 -----HHOIAKIKLSKGRVPLVGPEEVOFRLOALMECDTRKPKRLALSVYSOKMD 235
DB 594 FPNNEDELEIOGKLPDPVKEGCAPE-----WPNVEKLIKOCLKENQERTSAQVFDILNS 649
OY 236 PTFATFWELOCCGKOT---AFFSSOGQETVYVFDG---KEESRNTVAVTEKGLMEV- 287
DB 650 AELVCLTRILLPKNVIVECVATHHNSRNASIMWGCCHTRGQLSFLDLNTEGTSSEV 709
OY 288 -ORMCPCGMKVSQLOVORSIM--TATEDOKIYITTLKGMCPMLTPPOALDTPAVVTCF 343
DB 710 ADSRLICLAL---VHLPEKESWIVSGTOSGFLVINTEDGKRHTLEKMTDS---VTCL 763
OY 344 LAVPYIK--KNSYLVLAGLADGLVAVF 368

DB 764 YCNSFSKSKOKNPLVGTADGKLAIF 790

RESULT 12
AAG22172
ID AAG22172 standard; Protein; 369 AA.
XX AAG22172;
AC
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24999.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134286.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135144.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138840.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

```

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 8.2%; Score 315; DB 21; Length 369;
 Best Local Similarity 35.5%; Pred. No. 1.7e-19;
 Matches 82; Conservative 44; Mismatches 73; Indels 32; Gaps 9;

```

QY 16 EFRDASMLHLOHPCIALIGISHP-CPFALELAPLSSINWYLSNADSSIFPLGH 73
DB 133 QFOEVSMLALKHNHYRFGACIKPVCIVTEYAKGGSVROPLFR--KQNRVPLKL 190
QY 74 MLTKRIAYQIASGLAYLHKNNIIFCDLKSNDILWMSLDVKEHINKILSDYGISROSFH-E 132
DB 191 AVMQ--ALDVARGMAYVERNFHHDLSKNDLLI-SAD-----RSKINDPVARIEVQTE 243
QY 133 GALTVEGIPGIGAPDIRIYIDEKVDMSFGWLVYELLSGQRPALGHQLOIA-KLISK 191

```

```

DB 244 GMTPEGTVMKMAPEMIQRPYTKVDVYSFGIVLMEILITGLPPONNTRAVQAARAVNR 303
QY 192 GIR-----PVLGQPEVVOFRRLQALMECWDITKEKRPALASVSOMK 234
DB 304 GVRPTVPADCLPLVGE-----IWTRCMDADPEVRPCFAEIVILLE 343

RESULT 13
ID AAG22171 standard; Protein; 374 AA.
XX
AC AAG22171;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24998.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132408.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134216.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134378.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138340.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

```

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147028.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147943.
PR 09-AUG-1999; 99US-0147945.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

8 28; Score 315; DB 21; Length 374;
Best Local Similarity 35.58; Pred. No. 1.7e-19;

Matches 82; Conservative 44; Mismatches 73; Indels 32; Gaps 9;

QY 16 EFROENMHALQHCIVATIGISIRPL--CEALELAPLSSINTVLSNARDSSFIPLGH 73
DB 138 QFOEVSMLAFIKRPNIVRFIGACIKRPMWVCIVTEYKAGGSVRQFLTK--RONNAVPLKL 195

QY	16	EFRRASALNHLAHPFCIVALLGISIHPL--CFALDELAPLSLNTVLSENNRDSFFPLGH	73
Db	176	QFOEVSMTLAFELKRPNIYVRFGLACIKRPMWCITYETFAKGSVRFELK--RQNRVPLKL	233
QY	74	MLTOKIAVOIASGLAYLHKNNIFPCDKSDNIIWMSLDVKEHINIKLSDYGISRQSFH-E	132
Db	234	AVMG--ALDVAAGMAYVHERNFIRDLKSDNLLI-SAD---RSIKIADGVARIQVOTE	286
QY	133	GALVEGTPGTVQAPETIRPRIVYDEKVDMEFSGMYLYELLSSQRPALGHQQLA-KRLSK	191
Db	287	GMPTETGTGYRMMAPEMIOHREYTKVDVYSFGIYLMELITGLLFFQNMNTAVQAFAVYNR	346
QY	192	GIR-----PVLGOPEEVOFRLLQALMMECDPTKPEKRPALSYVSQMK	234
Db	347	GVREPTVPADCLPVLGE-----IMTRCMDADPEVRRCFAEIVNLE	386
RESULT	15		
AAG32053	ID	AAG32053 standard; Protein: 367 AA.	
XX	AAG32053:		
XX	17-OCT-2000	(first entry)	
DE	Arabidopsis thaliana	protein frgment SEQ ID NO: 38597.	
XX	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter		
KM	termination sequence.		
XX	Arabidopsis thaliana.		
OS	Ep1033405-A2.		
PN	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
XX	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126284.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	23-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132047.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134286.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135629.	
PR	25-MAY-1999;	99US-0136021.	

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147792.
PR 05-AUG-1999; 99US-0147760.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

RESULT 4
 LOCUS BC005408 3509 bp mRNA 11near PRI 12-JUL-2001
 DEFINITION Homo sapiens, similar to hypothetical protein FLJ23119, clone
 ACCSSION MGC:5630 IMAGE:3463078, mRNA, complete cds.
 VERSION BC005408.1 GI:13529310
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3509)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 REMARK
 COMMENT Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ARCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL) Genome
 DNA Sequencing by: Sequencing Group at the Stanford Human
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mdgpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
 Series: IRAC Plate: 4 Row: n Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 Location/Qualifiers
 1. 3509
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:5630 IMAGE:3463078"
 /tissue_type="Cervix, Carcinoma"
 /clone_id="NIH_MGC_12"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1317..1982
 /codon_start=1
 /product="Similar to hypothetical protein FLJ23119"
 /protein_id="AAH05408.1"
 /db_xref="GI:13529311"
 /translation="MYSTRYQLCARFCYSPRLRMPVRLPDIETEPASHTANK
 VPEBDSIADVSIMSEELGTOILTHQESLDYCKMSYSSPPQAAARSPSLSSPA
 SSSSVPTSDKEDMDLHMPGAASDRSHDTPMDGETFSQHLQAVILAVRDLIWP
 RRGCDVIVIGLEKDESEAOGRVIAVLKARELTPHGIPIVSVKVCNMGMPVDMVYMA
 AVM"
 BASE COUNT 810 a 994 c 953 g 752 t
 ORIGIN
 Query Match 87.7%; Score 1915.4; DB 9; Length 3509;
 Best Local Similarity 93.3%; Pred. No. 0;
 Matches 2074; Conservative 0; Mismatches 1; Indels 148; Gaps 1;
 QY 110 GATAGAGATCAACCCGCTCTGCGCCCTGAGCTCGCCGCTGACAGCCTCAACA 169
 DB 1 GATAGAGATCAACCCGCTCTGCGCCCTGAGCTCGCCGCTGACAGCCTCAACA 60
 QY 170 CCGTCTGTCGAGAGAGAGAGATCTCTCTTATACCCCTGGAGACATGCTACCC 229
 DB 61 CCGTCTGTCGAGAGAGAGAGATCTCTCTTATACCCCTGGAGACATGCTACCC 120
 QY 230 AAAAAATAGCTACAGATCGCTCGGGGCTGCTACCTGCACAAAGAAAATCATCT 289

DB 121 AAAAAATAGCTACAGATCGCTCGGGCTGCTACCTGCAACAAAGAAAATCATCT 180
 QY 290 TCTGTGACCTGAGAGTGGAGACACATCTGTGTGTGTCCTTGAAGTCAAGACATCA 349
 DB 181 TCTGTGACCTGAGAGTGGAGACACATCTGTGTGTGTCCTTGAAGTCAAGACATCA 240
 QY 350 ACATCAAGCTATCTGACTACGGGATTTCCAGGAGATCATTCATGAGGGCCCTAGCCG 409
 DB 241 ACATCAAGCTATCTGACTACGGGATTTCCAGGAGATCATTCATGAGGGCCCTAGCCG 300
 QY 410 TGAGAGGACTCTGCTGCTACAGGCCCCAGAGATCAGGCTCGATGTATGATAGA 469
 DB 301 TGAGAGGACTCTGCTGCTACAGGCCCCAGAGATCAGGCTCGATGTATGATAGA 360
 QY 470 AGGTAGATATGTTCTCTATGAGATGAGTGTCTACGAGTTGCTGTCAGACAGCCCTG 529
 DB 361 AGGTAGATATGTTCTCTATGAGATGAGTGTCTACGAGTTGCTGTCAGACAGCCCTG 420
 QY 530 CACTGGGCGCACACACAGCTCCAGATTCGCAAGAGCTGTCCAGGGGATCCGCCGGTTC 589
 DB 421 CACTGGGCGCACACACAGCTCCAGATTCGCAAGAGCTGTCCAGGGGATCCGCCGGTTC 480
 QY 590 TGGGCGAGCGGAGAGAGTGAAGTCCGGGAGCTGACAGGCGCTCATGATGAGAGTGGG 649
 DB 481 TGGGCGAGCGGAGAGAGTGAAGTCCGGGAGCTGACAGGCGCTCATGATGAGAGTGGG 540
 QY 650 ACATAAGCCAG 709
 DB 541 ACATAAGCCAG 600
 QY 710 CTTTGGCAGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 769
 DB 601 CTTTGGCAGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 770 AGGCGCAGAGAGTACAGGAGT 829
 DB 661 AGGCGCAGAGAGTACAGGAGT 720
 QY 830 TGTGACACACAG 889
 DB 721 TGTGACACACAG 780
 QY 890 TGAAGTGCACCTCAGCTCAGTCCAGAGATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 949
 DB 781 TGAAGTGCACCTCAGCTCAGTCCAGAGATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 QY 950 ACATCTACACCTCAAGGGGATGTGGCCCTTAACACACCCCAAGGCTTGATATCT 1009
 DB 841 ACATCTACACCTCAAGGGGATGTGGCCCTTAACACACCCCAAGGCTTGATATCT 900
 QY 1010 CAGCTGTGCTACCTGCTCTTGGCCGCTGCTGTTATTTAAAGAAATTCCTACCTGTCT 1069
 DB 901 CAGCTGTGCTACCTGCTCTTGGCCGCTGCTGTTATTTAAAGAAATTCCTACCTGTCT 960
 QY 1070 TAGCGGCGCTCGCGATGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1129
 DB 961 TAGCGGCGCTCGCGATGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
 QY 1130 ACAGTGTCTCTACCTGT 1189
 DB 1021 ACAGTGTCTCTACCTGT 1080
 QY 1190 AAGAGCAGAGGAGAACCCCTACCCAGTGAAGAGCCATGAGGTGTGTCAACAGCGGCTGTG 1249
 DB 1081 AAGAGCAGAGGAGAACCCCTACCCAGTGAAGAGCCATGAGGTGTGTCAACAGCGGCTGTG 1140
 QY 1250 AGGTGTGTACAGCATGGCGGCGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1309
 DB 1141 AGGTGTGTACAGCATGGCGGCGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
 QY 1310 GCAGGCGGTGAGCGCTACAGTGGCCCGCTCGATGATGATGATGATGATGATGATGATGATGAT 1369


```

Db 1249 TTCTCTTATGAAATGATGCTCTTACAGATGCTGTCAGACACGCGCTCTGCACTGGGCGAC 1308
QY 541 CACCAGCTCAGATATGCGCAGAAAGCTGTCCAAAGGCGATCCGCGGCTCTGAGGAGCGG 600
Db 1409 CACAGAGCTCAGATATGCGCAGAAAGCTGTCCAAAGGCGATCCGCGGCTCTGAGGAGCGG 1368
QY 601 GAGAAAGCTCAGATATGCGGCGACTGCGAGCGCTCATGATGAGTGTGTGGAGACTAAGCA 660
Db 1369 GAGAAAGCTCAGATATGCGGCGACTGCGAGCGCTCATGATGAGTGTGTGGAGACTAAGCA 1428
QY 661 GAGAAAGCTCAGATATGCGGCGACTGCGAGCGCTCATGATGAGTGTGTGGAGACTAAGCA 720
Db 1429 GAGAAAGCTCAGATATGCGGCGACTGCGAGCGCTCATGATGAGTGTGTGGAGACTAAGCA 1488
QY 721 TTCAATGATGATGCTGTGTGTGGAGAGACAGCCTCTTCTCATCCAGGCGCGAGAG 780
Db 1489 TTCAATGATGATGCTGTGTGTGGAGAGACAGCCTCTTCTCATCCAGGCGCGAGAG 1548
QY 781 TACACCTGTGTGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 1549 TACACCTGTGTGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1608
QY 841 GAGAAAGCTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 1609 GAGAAAGCTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1668
QY 901 CTCAGAGCTCAGAGATGCTGTGAGACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 1669 CTCAGAGCTCAGAGATGCTGTGAGACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1728
QY 961 CTCAGAGCTCAGAGATGCTGTGAGACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 1729 CTCAGAGCTCAGAGATGCTGTGAGACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1788
QY 1021 ACCTGCTCTTGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1789 ACCTGCTCTTGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1848
QY 1081 GCCGATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 1849 GCCGATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1908
QY 1141 TACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 1909 TACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1968
QY 1201 CAGAAAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1969 CAGAAAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2028
QY 1261 AGCAATGGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 2029 AGCAATGGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2088
QY 1321 GAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 2089 GAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2148
QY 1381 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 2149 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2208
QY 1441 ACCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 2209 ACCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2268
QY 1501 CCGGAGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
Db 2269 CCGGAGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2328
QY 1561 CCTGAGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620

```

```

Db 2329 CCTGAGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2388
QY 1621 ATCTGATGCACAGAGATATCTACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
Db 2389 ATCTGATGCACAGAGATATCTACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2448
QY 1681 CCACCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
Db 2449 CCACCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2508
QY 1741 AGTGTGCTCTTCTCAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 2509 AGTGTGCTCTTCTCAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2568
QY 1801 TCCGACAGCTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 2569 TCCGACAGCTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2628
QY 1861 CAGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
Db 2629 CAGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2688
QY 1921 GTATGCTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
Db 2689 GTATGCTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2748
QY 1981 TTTAAAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
Db 2749 TTTAAAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2808
QY 2041 GACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
Db 2809 GACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2868
QY 2101 GCGTGGGCGCGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
Db 2869 GCGTGGGCGCGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2928
QY 2161 GAGGCTGCTGCTGCGAAGAGAG 2184
Db 2929 GAGGCTGCTGCTGCGAAGAGAG 2952

RESULT 3
AX250161 6045 bp DNA linear PAT 28-SEP-2001
LOCUS AX250161
DEFINITION Sequence 6 from Patent WO016594.
ACCESSION AX250161
VERSION AX250161.1 GI:15864536
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 6045)
AUTHORS Plowman, G.D., Whyte, D., Manning, G., Sudarsanam, S. and Martinez, R.
TITLE Human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 016594-A 6 13-SEP-2001;
Sugen, Inc. (US)
FEATURES
source 1..6045
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1359 a 1779 c 1707 g 1200 t
ORIGIN
Query Match 88.7%; Score 1936.8; DB 6; Length 6045;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1;
QY 1 ATGCTGAGGACCTCGGGGCGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
Db 3976 ATGCTGAGGACCTCGGGGCGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4035

```

Db 301 CATGATGAGTGTCTGGACACTAAGCCAGAGAAGGACCGCTGGCCCTGTGGTGTGAG 360
 QY 693 CCAAGATGAAGAGACCCGACACTTTGGCCACTCATATGTAAGTGTGTGGAGACAGAC 752
 Db 361 CCAGATGAAGAGACCCGACACTTTGGCCACTCATATGTAAGTGTGTGGAGACAGAC 420
 QY 753 AGCGTCTCTATCCAGAGGAGAGTACACCGTGGTGTGGAGTGAAGAGAGA 812
 Db 421 AGCGTCTCTATCCAGAGGAGAGTACACCGTGGTGTGGAGTGAAGAGAGA 480
 QY 813 GTCCAGGAACCTACAGGTGTGTGAACACAGAGAAGGCGCTCATGAGAGTGCAGAGATGTG 872
 Db 481 GTCCAGGAACCTACAGGTGTGTGAACACAGAGAAGGCGCTCATGAGAGTGCAGAGATGTG 540
 QY 873 GTGCGCTGGAGTGAAGGTGAGTCCAGCTCAGTCCAGTCCAGAGATCCCTGTGAACAGCCAC 932
 Db 541 GTGCGCTGGAGTGAAGGTGAGTCCAGCTCAGTCCAGAGATCCCTGTGAACAGCCAC 600
 QY 933 CGAGGACCGAAGAAATCTATATCTACACCTCAAGGCGATGTGCCCTTAAACACAGCCCA 992
 Db 601 CGAGGACCGAAGAAATCTATATCTACACCTCAAGGCGATGTGCCCTTAAACACAGCCCA 660
 QY 993 ACAGGCGTTGATCTACTCAGCTGTGCTCAGTCTGTGGCGGCTGTATTATTTAAAA 1052
 Db 661 ACAGGCGTTGATCTACTCAGCTGTGCTCAGTCTGTGGCGGCTGTATTATTTAAAA 720
 QY 1053 GAATTCCTACCTGCTGTGTAGCGGGCTCGCGCATGGGCTGTGTGGTGTGTTCCTGGT 1112
 Db 721 GAATTCCTACCTGCTGTGTAGCGGGCTCGCGCATGGGCTGTGTGGTGTGTTCCTGGT 780
 QY 1113 GGGGGGACCCCAAGAGAGAGTGTCTCTACTGTGTCTACACACAGCCCAAGGCTCCAA 1172
 Db 781 GGGGGGACCCCAAGAGAGAGTGTCTCTACTGTGTCTACACACAGCCCAAGGCTCCAA 840
 QY 1173 GTTTCAGCATCGGGAGTGAAGAGCAGCAGCGAGAAACCCCTACAGTAAGGAGCATGGAGGT 1232
 Db 841 GTTTCAGCATCGGGAGTGAAGAGCAGCAGCGAGAAACCCCTACAGTAAGGAGCATGGAGGT 900
 QY 1233 GGTCAACAGCGGCTGTGAGGTGTGTGTACAGCAATGGGCGGCTCTGTGTATCGACTG 1292
 Db 901 GGTCAACAGCGGCTGTGAGGT 960
 QY 1293 TGGCTCCCTGGAGATCTGTAGGCGGT 1352
 Db 961 TGGCTCCCTGGAGATCTGTAGGCGGT 1020
 QY 1353 AGT 1412
 Db 1021 AGT 1080
 QY 1413 CAATCTCTTGT 1472
 Db 1081 CAATCTCTTGT 1140
 QY 1473 GGTGCCCAAGCCCTCAGGAGATGT 1532
 Db 1141 GGTGCCCAAGCCCTCAGGAGATGT 1200
 QY 1533 AGCCAGCCACAGCGGCGCAACCCAAAGGT 1592
 Db 1201 AGCCAGCCACAGCGGCGCAACCCAAAGGT 1260
 QY 1593 CATGTACAGTGAAGAGT 1652
 Db 1261 CATGTACAGTGAAGAGT 1320
 QY 1653 CGGTGTCATGTCTCTCTATCTCTCATCTCCAGCCCGCGAGGT 1712
 Db 1321 CGGTGTCATGTCTCTCTATCTCTCATCTCCAGCCCGCGAGGT 1380
 QY 1713 CCTCCCAAGCTCCCAAGAGT 1772
 Db 1381 CCTCCCAAGCTCCCAAGAGT 1440

QY 1773 AGACATGCTACATAGCCCGGT 1832
 Db 1441 AGACATGCTACATAGCCCGGT 1500
 QY 1833 GGAGGGGAGAGCTTCAGGACACCTGAGGCGGTGAAGATCTCTGGCGGTGAGAGACT 1892
 Db 1501 GGAGGGGAGAGCTTCAGGACACCTGAGGCGGTGAAGATCTCTGGCGGTGAGAGACT 1560
 QY 1893 CATTTGGGTCCCAAGCGCGGT 1952
 Db 1561 CATTTGGGTCCCAAGCGCGGT 1619
 QY 1953 AGCCAGCGGGGCGGAGTCAATGTCCTTTAAAGCCGAGAGT 2012
 Db 1620 AGCCAGCGGGGCGGAGTCAATGTCCTTTAAAGCCGAGAGT 1679
 QY 2013 GCTGT 2072
 Db 1680 GCTGT 1739
 QY 2073 CACAGAGT 2132
 Db 1740 CACAGAGT 1799
 QY 2133 CCAAGTCTACAGAGAGT 2184
 Db 1800 CCAAGTCTACAGAGAGT 1851

RESULT 6
 AK094556
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ37237 fls, clone BRAMY2002638.
 ACCESSION AK094556
 VERSION AK094556.1 GI:21753640
 KEYWORDS Oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens amygdala cDNA to mRNA, clone 11b.BRAMY2
 clone:BRAMY2002638.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
 Furukumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
 Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
 Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
 Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
 Wagaitsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahara,K., Masuho,Y., Nagai,K. and Isegai,T.
 NEDO human cDNA sequencing project
 Unpublished

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NEDO human cDNA sequencing project
 2 (bases 1 to 2790)
 Isegai,T. and Yamamoto,J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isegai, FLJ Project(HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomicehri.co.jp, tel:81-438-52-3975, fax:81-438-52-3986)
 Economy, Trade and Industry of Japan: cDNA full insert sequencing;
 Research Association for Biotechnology (RAB) (supported by Japan
 Key Technology Center etc.); 5'- & 3'- and one pass sequencing; RAB,
 HRI, and Biotechnology Center. National Institute of Technology and
 Evaluation; clone selection for full insert sequencing; HRI and
 RAB; annotation; HRI and RAB.

FEATURES
 SOURCE
 1. 2790
 Location/Qualifiers
 /organism: "Homo sapiens"
 /db_xref: "taxon:9606"
 /clone: "BRAMY2002638"
 /tissue_type: "amygdala"

```

Db 1201 GCGAGCGGCTGTAAGCCCTACATGAGCCCTCCATGTTACGTACGTGCTGACAGCTCAG 1260
QY 1370 AGGCAAGAGGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1429
Db 1261 AGGCAAGAGGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1430 ACCAGCTGCAAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1489
Db 1321 ACCAGCTGCAAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1450 GGGAGATGTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1549
Db 1381 GGGAGATGTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1550 ACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1609
Db 1441 ACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1610 TGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1669
Db 1501 TGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1670 ACTCTCTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1729
Db 1561 ACTCTCTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1730 CAAGTTCTTCCAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1789
Db 1621 CAAGTTCTTCCAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
QY 1790 CCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1849
Db 1681 CCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
QY 1850 GCCAGACCTGAGAGGCGCTGAAGATCTGCGCTGAGAGAGAGAGAGAGAGAGAGAG 1909
Db 1741 GCCAGACCTGAGAGGCGCTGAAGATCTGCGCTGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1910 GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1969
Db 1801 GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1970 TCATTGCTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2008
Db 1861 TCATTGCTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 2009 ----- 2008
Db 1921 TGAAGGATGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
QY 2009 ----- 2008
Db 1981 GAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
QY 2009 ----- 2008
Db 2041 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
QY 2062 GAAAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2121
Db 2101 GAAAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 2122 GACATTTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2181
Db 2161 GACATTTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2182 AGG 2184
Db 2221 AGG 2223

```

RESULT 5
AK054662

```

LOCUS      AK054662      3112 bp      mRNA      linear      PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ30100 fis, clone BNGH41000104.
ACCESSION  AK054662
VERSION    AK054662.1  GI:16549250
KEYWORDS   Oligo cloning; fis (full insert sequence).
SOURCE     Homo sapiens neuroglioma cell_line:HB. cDNA to mRNA,
           clone_11b:BNGH41 clone:BNGH41000104.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1
AUTHORS   Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
            Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
            Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
            Sekine,M., Kikuchi,H., Kanda,K., Wagaetsuma,M., Murakawa,K.,
            Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A.,
            Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Masuno,Y., Nagai,K.
            and Isogai,T.
TITLE      NEDO human cDNA sequencing project
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 3112)
AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE      Direct Submission
JOURNAL    Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yano, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp; Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT    NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.
FEATURES   source
            1..3112
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="BNGH41000104"
               /cell_line="H4"
               /cell_type="neurogloma"
               /clone_11b="BNGH41"
               /note="cloning vector: pME18SF13"
BASE COUNT 707 a 872 c 862 g 671 t
ORIGIN
Query Match      84.2%; Score 1840; DB 9; Length 3112;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 333 CGTCAGAGAGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 392
Db 1 CGTCAGAGAGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 60
QY 393 TGAAGGCGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
Db 61 TGAAGGCGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 453 CATTGATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 512
Db 121 CATTGATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 513 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 572
Db 181 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 573 GGGCATCCGCGGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 632
Db 241 GGGCATCCGCGGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 633 CATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692

```

ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0656.0370002/SLE/LBB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3561
US-08-665-574C-12

Query Match 2.1%; Score 46.4; DB 3; Length 3561;
Best Local Similarity 47.9%; Pred. No. 0.076;
Matches 134; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 236 TAGCCACAGATCGCCCTGGGCGCTGCTACCTGACAAAGAAAACATCATCTTGTG 295
DB 3008 TCGCCACGAGATCGCGAGGGCGCTATCTGACGCGACGATCACCAG 3067
QY 296 ACCCTGAAGTCGACACATCTGTGTGCTGCTGACGTCACAGACATCAATCA 355
DB 3068 ACCTGACCGCGGCAACGTCGTGTGACACAGACGCTGTGCAAGATCGGGACTTG 3127
QY 356 ACCTATCTACTACGAGATTTGAGGCACTCATTCATAGAGGCGCCCTAGCGGTGAGG 415
DB 3128 GCTTACACCAAGCGCGCCGACAGACAGATACCGCGCGCGAGAGTGGGAG 415
QY 416 GCACCTCTGCTACGAGCGCCGACGATCAGGCTGATATATGATGAGAAAGTAG 475
DB 3188 GCCCGCTGTCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3247
QY 476 ATATGTTCTCTATGGAATGTGCTCTACAGAGTTGCTGTC 515
DB 3248 ATGTCTGTCTTCTGCGGGTGACCTGTATGAGCTGCTGAC 3287

RESULT 12
US-08-946-994-12
Sequence 12, Application US/08946994
Patent No. 6210654
GENERAL INFORMATION:
APPLICANT: Inle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Wiltuhn, Bruce A.
APPLICANT: Quelle, Frederick W.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
TITLE OF INVENTION: Transduction
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,994
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574

FILING DATE: 18-JUN-1996
APPLICATION NUMBER: 08/282,012
FILING DATE: 29-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/097,997
FILING DATE: 29-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/118,968
FILING DATE: 09-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0370002/SLE/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3561
US-08-946-994-12

Query Match 2.1%; Score 46.4; DB 4; Length 3561;
Best Local Similarity 47.9%; Pred. No. 0.076;
Matches 134; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 236 TAGCCACAGATCGCCCTGGGCGCTGCTACCTGACAAAGAAAACATCATCTTGTG 295
DB 3008 TCGCCACGAGATCGCGAGGGCGCTATCTGACGCGACGATCACCAG 3067
QY 296 ACCCTGAAGTCGACACATCTGTGTGCTGCTGACGTCACAGACATCAATCA 355
DB 3068 ACCTGACCGCGGCAACGTCGTGTGACACAGACGCTGTGCAAGATCGGGACTTG 3127
QY 356 ACCTATCTACTACGAGATTTGAGGCACTCATTCATAGAGGCGCCCTAGCGGTGAGG 415
DB 3128 GCTTACACCAAGCGCGCCGACGATCAGGCTGATATATGATGAGAAAGTAG 475
QY 416 GCACCTCTGCTACGAGCGCCGACGATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3247
QY 476 ATATGTTCTCTATGGAATGTGCTCTACAGAGTTGCTGTC 515
DB 3248 ATGTCTGTCTTCTGCGGGTGACCTGTATGAGCTGCTGAC 3287

RESULT 13
US-08-755-728-1
Sequence 1, Application US/08755728
Patent No. 5962312
GENERAL INFORMATION:
APPLICANT: Ploman, Gregory
APPLICANT: Mossie, Kevin
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage

Query Match	2.2%	Score 47.4	DB 4	Length 2543
Best Local Similarity	51.7%	Pred. No. 0.039		
Matches 170; Conservative	0	Mismatches 141	Incls 18	Gaps 2

RESULT 10
US-08-097-997A-12
; Sequence 12, Application US/08097997A

1 Patent No. 5728536
2
3 GENERAL INFORMATION:
4 APPLICANT: Ihle, James N.
5 APPLICANT: Silvenoinen, Ollie
6 APPLICANT: Witthuhn, Bruce A.
7 APPLICANT: Quelle, Frederick W.
8 TITLE OF INVENTION: Jak Kinases and Regulation of
9 TITLE OF INVENTION: Transduction
10 NUMBER OF SEQUENCES: 14
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
13 STREET: 1100 New York Avenue, Suite 600
14 CITY: Washington
15 STATE: D.C.
16
17 COUNTRY: U.S.A.
18 ZIP: 20005-3934
19
20 COMPUTER READABLE FORM:
21
22 MEDIUM TYPE: Floppy disk
23
24 COMPUTER: IBM PC compatible
25 OPERATING SYSTEM: PC-DOS/MS-DOS
26 SOFTWARE: patentln Release #1.0, Version #1.25
27
28 CURRENT APPLICATION DATA:
29 APPLICATION NUMBER: US/08/097,997A
30 FILING DATE: 29-JULY-1994
31
32 CLASSIFICATION: 435
33
34 ATTORNEY/AGENT INFORMATION:
35
36 NAME: Fox, Samuel L.
37 REGISTRATION NUMBER: 30,353
38 REFERENCE/DOCKET NUMBER: 0656, 0370000/SLE/GKT
39
40 TELECOMMUNICATION INFORMATION:
41
42 TELEPHONE: (202) 371-2600
43

```

Query Match      2.1%      Score 46.4; DB 1; length 3561;
Best Local Similarity 47.9%      Pred. No. 0.076;
Matches 134; Conservative. 0; Mismatches 146; Indels 0; Gaps 0;

```

RESULT 11
US-08-665-574C-12
; Sequence 12, Application US/08665574C

Patent No. 616595
GENERAL INFORMATION:
APPLICANT: Inle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Miltihm, Bruce A.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine
TITLE OF INVENTION: Signal Transduction
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Stene, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574C
FILING DATE: 18-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/282,012
FILING DATE: 29-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/097,997
FILING DATE: 29-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/118,968
FILING DATE: 09-SEP-1993

QY 245 AGATGGCTGGGCTGGGCGCTTACTCTCAACAAGAAAACTCATCTTGTGTGACCTGAAGT 304
 Db 725 AGATCAGCGCTGGGCGCTTGGGCACTCTCACTCCAGGGCATCTATCTACCGGAGACTCAAGC 784
 QY 305 CGGACAACATTTGTGTGTGTGCTCCCTTGAGCTCAAGGACCAATCAATCAATCAAGTATGTG 364
 Db 785 CCGGAACAATCATATGCTTCAGGACGCAAGGGCC-----ACATCAAAACGAGCCG 829
 QY 365 ACTACGGATTTTGGAGGACATCTATTCATGAGGGCGCCCTAGG-----CGTGAGGGCA 418
 Db 830 ACTTTGACCTCGAAGGATCTATCTCATGAGGGCGCGCTCACTCAACACTTCTGGGCA 889
 QY 419 CTCTGGCTACGAGGCGCCCAAGAGATAGGCGCTTGCATTTGTATATGATGAGAAGTATATA 478
 Db 890 CCATTTGATCAATGCGCCCTCGAATATTTCTGTGGCGATGTGGCAACCGGGGTGTGGACT 949
 QY 479 TGTTCCTCTATGGAATGATGATCTCTACAGATTTGCTGTCAAGAGCGCCCTGCATGGGCC 538
 Db 950 GGTGAGACCTGGGGGCGCTGATGTATGACATGCTCACTGATGCGCGCGCTTCAACCGAG 1008
 QY 539 ACCACAGCTCCAGATTTGCCAAGCAACCTGCTCAAGGGCA 577
 Db 1010 AGAATCGGAAGAAACCATGTGATTAAGTATCTCAGGGGCA 1048

```

RESULT 7
US-09-430-564-1
; Sequence 1, Application US/09430564
; Patent No. 6372467
; GENERAL INFORMATION:
; APPLICANT: John Blenis
; APPLICANT: Ray K. Lee-Fruman
; APPLICANT: Calvin J. Kuo
; TITLE OF INVENTION: P5456K AND P8556K GENES, PROTEINS,
; TITLE OF INVENTION: PRIMERS, PROBES, AND DETECTION METHODS
; FILE REFERENCE: 00246/506002
; CURRENT APPLICATION NUMBER: US/09/430,564
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,141
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-430-564-1

```

	Query Match	2.3%	Score 49.8:	DB 4:	Length 1732:
	Best Local Similarity	51.9%:	Pred. No. 0.0094:		
	Matches 176:	Conservative	0:	Mismatches 142:	Indels 21; Gaps 2.
OY	245	AGATGCGCCTGGGCCCTGTGCCTACTGACACAAGAAAAACATCATCTTCTGTGACTGTAAGT	304		
Db	565	AGATACCCCTGGCCCCCTGGGCCATCTTCACCTCCAGGGCATCTCTACCGGGGCACTTCACG	624		
OY	305	CGGACACATTCTGGTGTTGTTGCTTGAAGTCAGAAGACATCAACATCAAGCTATCTG	364		
Db	625	CGGAAACATATGTTGTCAGCACAGCGGGC-----ACATCAACTGACC	669		
OY	365	ACTACGGGATTTTCAGGCGATCTTCATGAGGGGCGCCCTAG-----CGTGAGGGCA	418		
Db	670	ACTTTGGACTCTGCAGAAGACTTATCCATGAGGGGTCGCGCTCACTCACACTTCTGGGGCA	729		
OY	419	CTTCGCGCIACCAGGCCCCAGAGATCAGGCCCTGCATTTATATGATGGAAGTAGATA	478		
Db	730	CCATTGAATACATGGCCCCCTGAGATTTCTGGTGGCAGATGGCCACAACCGGGCTGGAGCT	789		
OY	479	TGTTCTCGATGGATGAGTGCCTCTGACGATTTCTGTCAAGAACAGGGCCCTGCACGTGGCC	538		
Db	790	GCTGAGGCTGGGGGCCCTGATGATGCAGATCTACTGAGATCGCGCCCTTACCGCAG	849		

QY 539 ACCACGAGCTCCAGATTGCCAAGAAGCTGTCCAAGGGCA 577
| | | | | | | | | |
Db 850 AGAACCGGAGAAAAACCATGGATAAGATCATCAGGGGCA 888

```

RESULT 8
US-09-429-322-3
Sequence 3, Application US/09429322A
Patient No. 6190869
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF PROTEIN KINASE C-THETA
TITLE OF INVENTION: EXPRESSION
FILE REFERENCE: RTS-0100
CURRENT APPLICATION NUMBER: US/09/429,322A
CURRENT FILING DATE: 1999-10-26
NUMBER OF SEQ. ID NOS: 89
SEQ ID NO 3
LENGTH: 2754
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (95)..(2215)
US-09-429-322-3

```

[illegible]

RESULT 9
 US-09-467-082-3
 : Sequence 3, Application US/09467082
 : GENERAL INFORMATION:
 : APPLICANT: Brett P. Monia
 : APPLICANT: Lex M. Cowsett
 : TITLE OF INVENTION: ANTISENSE MODULATION OF PKA CATALYTIC SUBUNIT C-ALPHA EXPRESSION
 : FILE REFERENCE: RFS-0088
 : CURRENT APPLICATION NUMBER: US/09/467,082
 : CURRENT FILING DATE: 1999-12-17
 : NUMBER OF SEQ ID NOS: 49
 : SEQ ID NO 3
 : LENGTH: 2549

ATTORNEY/AGENT INFORMATION:

NAME: Saxe, Stephen A.
REGISTRATION NUMBER: 38,609
REFERENCE/DOCKET NUMBER: 2323-120A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-624-1589
TELEFAX: 202-783-6031

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:
LENGTH: 3262 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 96..2036
US-08-678-039A-41

Query Match 2.4%; Score 52.8; DB 2; Length 3262;

Best Local Similarity 52.2%; Pred. No. 0.0022;

Matches 117; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

DB 335 TCAGGAGCAGCATCAACATCAAGTATCTGACTACGGGATTTGAGGCGAGTATTCATG 394

DB 1546 TCATGCTGAGCAGAGACTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCG 1605

DB 395 AGGGGCGCTAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAG 454

DB 1606 AAGAGCGCTACACCGGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAG 1665

DB 453 TTGTATATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 514

DB 1666 GCAGCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1725

DB 515 CAGGAGCGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCT 558

DB 1726 GCGGCGTGAACGACGACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1769

RESULT 5

US-09-800-960-1

Sequence 1, Application US/09800960

Patent No. 6387677

GENERAL INFORMATION:

APPLICANT: YE, Jane et al.

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

FILE REFERENCE: CLO01158

CURRENT APPLICATION NUMBER: US/09/800,960

CURRENT FILING DATE: 2001-03-08

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 2061

TYPE: DNA

ORGANISM: Human

US-09-800-960-1

Query Match 2.4%; Score 51.4; DB 4; Length 2061;

Best Local Similarity 47.4%; Pred. No. 0.0041;

Matches 154; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

DB 242 ACCAGATCCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

DB 465 ATCAGATTTCTGAGAGTGTATACCATCCACGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 524

DB 302 AGTCGAGCAACATTCGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361

DB 535 AGCTGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 584

DB 362 CTGACTAGGAGATTTGAGGCGCTATTCATGAGGCGGCTGAGGCGCTGAGGCGCTGAGGCGCT 421

DB 585 TTGGCTAGCCATGCAAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 644

DB 422 CTGGCTACAGGCGGCGGAGATGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481

DB 645 CAGGTACTTGTCCCTGAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 704

DB 482 TCTCTATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541

DB 705 GGGCTGCGGGGTATCTGCTGATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764

DB 542 ACCAGCTCCAGATTCGCAAGAGCT 566

DB 765 ATCAGCACAGCTGTATACGACGAT 789

RESULT 6

US-08-966-316-10

Sequence 10, Application US/08966316

Patent No. 5932445

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Au-Young, Janice

APPLICANT: Reddy, Roopa

APPLICANT: Murthy, Lynn E.

TITLE OF INVENTION: SIGNAL PEPTIDE - CONTAINING PROTEINS

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/966,316

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0424 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1637 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: SKINBIT01

CLONE: 1863688

US-08-966-316-10

Query Match 2.3%; Score 49.8; DB 2; Length 1637;

Best Local Similarity 51.9%; Pred. No. 0.0093;

Matches 176; Conservative 0; Mismatches 142; Indels 21; Gaps 2;

D _b	366	GACACTAAGCCAGAGAAAGCGACCACTGGCCCTGTGGGTGAAGCCAGATGAAGGACC	425
Q _y	708	GACATTTGCCACCTTCATGATGAACCTGTGCTGTGGGAACAGACACCTTTCTC-TCAT	766
D _b	426	GACATTTGCCACCTTCATGATGAACCTGTGCTGTGGGAACAGACACCTTTCTTCAT	485
Q _y	767	CCGAGGGCCAGAGATACACCGT-AGTCTTTTGGGATGGAAA	806
D _b	486	NCGAGGGCCAGAGATACACCTGTGGGTCTTTTGGGATGGAAA	526

RESULT 2

```

US-09-948-802-10
? Sequence 10, Application US/09948802
? Patent No. 6465232
? GENERAL INFORMATION:
? APPLICANT: ROBINSON, KEITH E
? TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
? TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
? FILE REFERENCE: NMI-090
? CURRENT APPLICATION NUMBER: US/09/948,802
? CURRENT FILING DATE: 2001-09-07
? PRIOR APPLICATION NUMBER: 09/387,212
? PRIOR FILING DATE: 1999-08-31
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 10
? LENGTH: 526
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURES:
? OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-948-802-10

```

Query Match	21.98;	Score 477.6;	DB 4;	Length 526;
Best Local Similarity	98.58;	Pred. No. 1.4e-104;		
Matches 513; Conservative	0;	Mismatches 5;	Indels 3	

QY	289	TTCTGTGACCTGAAGTGGAGCAACAATCTGATGGTGGCCCTTGACGCTCAAGAGGACACATC	34.8
Db	6	TCCCGTAGCCGGAAGTGGGACACACATTCCTGCTGGTGGTCCCTTGACGCTCAAGAGGACACATC	65
QY	349	AACATCAAGCTATCTGACATACGGGATTTTCGAGGCAAGTCATTTCCATGAGGAGGCCCTTAGGC	40.8
Db	66	AACATCAAGCTATCTGACATACGGGATTTTCGAGGCAAGTCATTTCCATGAGGAGGCCCTTAGGC	125
QY	409	GTGGAGGGCACTCCTGGCTACACAGGCCCCAGAGATCAAGGCCCTGGCATTTGATATGATGAG	46.8
Db	126	GTGGAGGGCACTCCTGGCTACACAGGCCCCAGAGATCAAGGCCCTGGCATTTGATATGATGAG	185
QY	469	AAGTAGATATGTTCTCCTATGAGAAATGGTGCCTACAGATTTGGCTGAGGACAGGCCCT	52.8
Db	186	AAGTAGATATGTTCTCCTATGAGAAATGGTGCCTACAGATTTGGCTGAGGACAGGCCCT	24.5
QY	529	GCACTGGGCCACACACAGCTCCAGATTGCCAAGAAAGCTGTCCAAAGGCGCATCCGCGGTT	58.8
Db	246	GCACTGGGCCACACACAGCTCCAGATTGCCAAGAAAGCTGTCCAAAGGCGCATCCGCGGTT	30.5
QY	589	CTGGGGCGACCCGGAGGAAGTGCACATTTCCGGCGACTGCGAGCGCTCATGATGAGTGTCTG	64.8
Db	306	CTGGGGCGACCCGGAGGAAGTGCACATTTCCGGCGACTGCGAGCGCTCATGATGAGTGTCTG	36.5
QY	649	GACACTAAGCCAGAGAGCGACGCGCTGGCCCTGTCCGGTGGTGTG-AGCCAAATGTAAGGACCC	70.7
Db	366	GACACTAAGCCAGAGAGCGACGCGCTGGCCCTGTCCGGTGGTGTGAGCCAGATGTAAGGACCC	42.5
QY	708	GACATTTGCCACCTTCATGTGTGAACCTGTGCTGGGAGACAGACAGCCCTTCTTC-TCAT	76.6
Db	426	GACATTTGCCACCTTCATGTGTGAACCTGTGCTGGGAGACAGACAGCCCTTCTTCTTCAT	48.5
QY	767	CCGACGGGCGAGAGTACACCGT-GGTGTTTTGGAGTGAATA	80.6

Db 486 NCCAGGGCCAGGAGTACACTGTGGGTGTTTGGGATGAAA 526

RESULT 3
5266464

Patent No. 5266464
 APPLICANT: HOUSEY, GERARD
 TITLE OF INVENTION: METHOD OF SCREENING FOR PROTEIN INHIBITORS
 AND ACTIVATORS
 NUMBER OF SEQUENCES: 3
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/392,073
 FILING DATE: 10-AUG-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 154,206
 FILING DATE: 10-FEB-1988
 SEQ ID NO.: 1
 LENGTH: 2599
 5266464-1

Query Match	2.4%	Score	52.8	DB	6	Length	2599
Best Local Similarity	52.1%	Pred. No.	0.0021				
Matches	148	Conservative	0	Mismatches	127	Indels	9
						Gaps	1

OY	245	AGATGGCTGGGCGTCCCTACTGTGCACAAGAAAAACATCATCTCTGTGGACCTGAAT	304
Db	1437	AGATTGCCATCGGCTTTTCTTCTTGCAAGACAAAGGCATCATTTTACCGTAGCCTBAAC	1496
OY	305	CGGACAACATTCGTGTGTGCTCCCTTGACGTCAAGGACGACATCAACATCAAGCTATCTG	364
Db	1497	TTGACACAGTGATGCTGATTCG-----AGGGGCAATCAAAAATGCGTAGCTTGG	1544
OY	365	ACTACGGGATTTGAGGCATCATTTCCATGAGGGGCCCTTAGGCCTGAGGGCACCTCTGG	424
Db	1548	GCAATGCTTAAGAATAATATCTGGGATGGGTGACAAACCAAGACATTCGTGTGGACCTCAG	1607
OY	425	GCTACACAGGCCCCAGAGATCAGGCGCTCGCATTTATATGATGAGAGTAGATATGTCT	484
Db	1608	ACTACATGCCCCAGATATGATTCATTATCAGCCCTAAGGAAGTCTGTGAGCTGTGG	1667
OY	485	CCTATGGAATGTGCTCTACAGAGTGTGTCACAGACAGCGCCT	528
Db	1668	CGTTGGAGTCTGCTGATGAATGTTGGCTGCGGACGACCT	1711

RESULT 4

US-08-678-039A-41
Sequence 41, Application US/08678039A
Patent No. 5858662
GENERAL INFORMATION:
APPLICANT: Keating, Mark T.
APPLICANT: Morris, Colleen A.
TITLE OF INVENTION: Diagnosis of Williams Syndrome and
TITLE OF INVENTION: Williams Syndrome Cognitive Profile by Analysis of the
TITLE OF INVENTION: Presence or Absence of a LIM-Kinase Gene
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
STREET: 555 Thirteenth Street, N.W., Suite 701 East
STREET: Tower
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/678,039A
FILING DATE: 10-Jul-1996
CLASSIFICATION: 435

QY 2177 AGAAGG 2184
DB 428 AGAAGG 435

RESULT 15
B0064084

LOCUS B0064084 1072 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6854302 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925776
5', mRNA sequence.

ACCESSION B0064084
VERSION B0064084.1 GI:19892410

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1072)

NIH-MGC <http://mgi.mcl.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

Tissue Procurement: Lou Staudt

CDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1C82098 Row: a Column: 09

High quality sequence stop: 638.

FEATURES
source

1..1072

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5925776"

/clone_lib="NIH_MGC_99"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGAG(G). Size-selected >500bp for average insert size

1.8kb. Library constructed by Ling Hong in the laboratory

of Gerald M. Rubin (University of California, Berkeley)

using ZAP-cDNA synthesis kit (Stratagene) and Superscript

II RT (Life Technologies). Note: this is a NIH_MGC

Library."

BASE COUNT 214 a 331 c 308 g 219 t

ORIGIN

Query Match 18.9%; Score 413; DB 14; Length 1072;

Best Local Similarity 100.0%; Pred. No. 1.3e-87;

Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1772 CAGACATGCTACATACGCGGCTGCTGCTCGACAGCTGAGCTGACCTGACCCCA 1831

DB 1 CAGACATGCTACATACGCGGCTGCTGCTCGACAGCTGAGCTGACCTGACCCCA 60

QY 1832 TGGACGGGAGACCTTACACGACCACTGACCGCTGAAGATCCTCGCGTCAGAGACC 1891

DB 61 TGGACGGGAGACCTTACACGACCACTGACCGCTGAAGATCCTCGCGTCAGAGACC 120

QY 1892 TCATTGGGTCCCGAGCGGCTGAGATGATGCTATTGGCCCTGGAGAGATCTTG 1951

DB 121 TCATTGGGTCCCGAGCGGCTGAGATGATGCTATTGGCCCTGGAGAGATCTTG 180

QY 1952 AAGCCAGGCGGCGGAGCTATGCGCTTAAAGCCGAGAGCTGATCCGATGGGG 2011

DB 181 AAGCCAGGCGGCGGAGCTATGCGCTTAAAGCCGAGAGCTGATCCGATGGGG 240

QY 2012 TCGTGTGATCTGCGGTGGTGGCAAGACATGTTGTGCACTTTGAAATGAAA 2071

DB 241 TCGTGTGATCTGCGGTGGTGGCAAGACATGTTGTGCACTTTGAAATGAAA 300
QY 2072 ACACAGAGTGTGCTGCTGCGGCTGTGAGAGGCTGGGGCCCGACAGAGTTCGACATTTTCT 2131
DB 301 ACACAGAGTGTGCTGCTGCGGCTGTGAGAGGCTGGGGCCCGACAGAGTTCGACATTTTCT 360
QY 2132 ACCAGTCTACAGAGAGCTGCGGCTGTGAGAGGCTTGCACCTGCAAGAGAGAG 2184
DB 361 ACCAGTCTACAGAGAGCTGCGGCTGTGAGAGGCTTGCACCTGCAAGAGAGAG 413

Search completed: April 15, 2003, 05:33:25
Job time : 2117 secs

QY	913	AGATCCCGTGGACAGCCACGAGAGACAGAAATCTACATCTACACCTCTCAAGGCATG	972
DB	481	AACCTGCGTGTGGAGCGCCACGAGGACCAAGATCTCATCTACAGCTCTCAAGGCATG	540
QY	973	TGCCCCCTAAACACACCCCAAC	994
DB	541	TGTCACATGGAACACACCCACG	562
RESULT 9			
LOCUS	BOJ35590	484 bp	mRNA
DEFINITION	PMO-MT0202-090501-013-h11 MT0202 Homo sapiens CDNA, mRNA sequence.		
ACCESSION	BOJ35590		
VERSION	BOJ35590.1	GI:20991150	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 484) Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., Geoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
MEDLINE	20202653		
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-7704922 Fax: +55-11-7707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO022-PMO-MT0202-090501-013-h11et3-2001-05-09et4-1) Seq primer: puc 18 forward High quality sequence start: 18 High quality sequence stop: 477. Location/Qualifiers 1..484 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="MT0202" /dex_stage="Adult" /note="Organ: marrow; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESYES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		
BASE COUNT	86 a	173 c	127 g
ORIGIN		98 t	
Query Match	20.8%	Score 454.8	DB 14
Best Local Similarity	39.2%	Pred. No. 1e-97	Length 484
Matches 478:	Conservative	0	Mismatches 2
		Inuels	2
		Gaps	2
QY	1276	CTCCTTGATCATGACATGTCGCTCCCTGAGATCTGACGCGCTGGACCTCTACATGGCTC	1335
DB	3	CTCCTTGATCATGACATGTCGCTCCCTGAGATCTGACGCGG-TGAGACCCCTACATGGCTC	6
QY	1336	CCCTCCATGGATGACATGTCGCTGGAGCTGTGAGGCGAGAGGGGAGAGAGTGGCTG	1395
DB	62	CCCTCCATGGATGACATGTCGCTGGAGCTGTGAGGCGAGAGGGGAGAGAGTGGCTG	121

OY	1396	TGCCGATGTACAAAGCCCAACTCTTGATGTACACTCACCACCTTACAGCTGTGT	1455
Dd	122	TGCCTGGATTGACAAGGCCAACAATCCTTTGGTGAATGTACACTCCACCACCTTACAGCTGTGT	181
OY	1456	GCCCGGTAATCTTCGTGCGGGGTGCCAGCCCCCTTCAGGGACATGTTTTCCGTGCGGCCCTTGG	1515
Dd	182	GCCCGGTAATCTTCGTGCGGGGTGCCAGCCCCCTTCATGACATATGTTTTCCGTGCGGCCCTTGG	241
OY	1516	GACACGGGAACCCCCGGGACGACGCCAACCGGCAACCCAAAAGTGCCTGAGGGGGACATCCC	1575
Dd	242	GACACGGGAACCCCCGGGACGACGCCAACCGGCAACCCAAAAGTGCCTGAGGGGGACATCCC	301
OY	1576	ATCCGGGACAGTGA-6CATCATGTATACAGTAGAGAGTGGGGACGACAGATCCTATGATCACCA	1634
Dd	302	ATCCGGGACAGTGAAGGCAATCATGTATACAGTAGAGAGTGGGGACGACAGATCCTATGATCACCA	361
OY	1635	GGAATCACTCACTCACTCACTATGCTCCATGTCCTCTTACTCTTCATCCATCCACACCCCGGACAGC	1694
Dd	362	GGAATCACTCACTCACTCACTATGCTCCATGTCCTCTTACTCTTCATCCATCCACACCCCGGACAGC	421
OY	1695	TGCCAGGTGCCCCCTTCAAAGCTCCCAAGCTCCCGACAGTCTTCCAGTGTGCTTCTC	1754
Dd	422	TGCCAGGTGCCCCCTTCAAAGCTCCCAAGCTCCCGACAGTCTTCCAGTGTGCTTCTC	481
OY	1755	CA 1756	
Dd	482	AA 483	
RESULT 10	B8649591	715 bp mRNA linear EST 26-OCT-2001	
LOCUS	B8649591		
DEFINITION	B8649591 RIKEN full-length enriched; 16 days embryo head Mus		
ACCESSION	MUSCULUS CDNA clone C130082A22.5, mRNA sequence.		
VERSION	B8649591		
KEYWORDS	B8649591.1 GI:16483846		
SOURCE	EST.		
ORGANISM	house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 715)		
REFERENCE	Arakawa,T., Carrinci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hoyama,K., Horii,F., Ishii,Y., Ito,M., Kawai,Y., Konno,H., Kouda , M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki, 'D., Shibata,K.A., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyata,T., Muramatsu,M. and Hayashizaki,Y.		
AUTHORS	RIKEN Mouse ESTs (Arakawa,T., et al. 2001) Unpublished (2001)		
TITLE	Contact: Yoshihide Hayashizaki —		
JOURNAL	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute		
COMMENT	The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsr.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carrinci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh , M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) waghi,K., Fujiwaga,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunoda ,S., Kawai,Y., Okazaki,Y., Muramatsu,M., Inoue,Y., Kirai,A. and Hayashizaki,Y.		
	RIKEN integrated sequence analysis (RISA) system. 384-Turnout sequencing pipeline with 384 multipipillary sequencer. Genome Res 10 (11), 1757-1771 (2000)		

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12268. row: C column: 22
High quality sequence stop: 686.
Location/Qualifiers

FEATURES

source
1. 996
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5552973"
/clone_lib="NH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NH-MGC library."
BASE COUNT 221 a 323 c 271 g 180 t 1 others
ORIGIN

Query Match 21.7%; Score 473.4; DB 13; Length 996;
Best Local Similarity 98.1%; Pred. No. 5e-102;
Matches 521; Conservative 0; Mismatches 6; Indels 4; Gaps 4;

OY 1 ATGCTGAGGACACCTGCGGCGCCACGATGCAAGAACTTCCGAGTCCGCGAGAG 60
DB 427 ATGCTGAGGACACCTGCGGCGCCACGATGCAAGAACTTCCGAGTCCGCGAGAG 486
OY 61 GCCAGCATGCTGCAGCGCGCTGCAGCACCCTGCATGTCGCGCTCATCGCATCAGCATC 120
DB 487 GCCAGCATGCTGCAGCGCGCTGCAGCACCCTGCATGTCGCGCTCATCGCATCAGCATC 546
OY 121 CACCGGCTGTGTTGGCCCTGAGAGCTCGCGCGCTCAGCAGCCTCAACACCGTGTCTCC 180
DB 547 CACCGGCTGTGTTGGCCCTGAGAGCTCGCGCGCTCAGCAGCCTCAACACCGTGTCTCC 606
OY 181 GAGAACGCGAGATGCTTCTTATACCCCTGGGACACATCTCAACCAAAAATAGCC 240
DB 607 GAGAACGCGAGATGCTTCTTATACCCCTGGGACACATCTCAACCAAAAATAGCC 666
OY 241 TACAGATGCGCTCGGCGCTGCAGTGCACAGAAAAACATCATCTTCTGTGACCTG 300
DB 667 TACAGATGCGCTCGGCGCTGCAGTGCACAGAAAAACATCATCTTCTGTGACCTG 726
OY 301 AAGTGGGACACAT-TCTGTGTGTTCTTCACTGCAAGAGACATCAATCAAGT 359
DB 727 AAGTGGGACACAT-TCTGTGTGTTCTTCACTGCAAGAGACATCAATCAAGT 786
OY 360 ATCTGACTACGGGATTTGAGGACAGATTCATGAGGCGCCCTAGAGGCTGAGGAGGAC 419
DB 787 ATCTGACTACGGGATTTGAGGACAGATTCATGAGGCGCCCTAGAGGCTGAGGAGGAC 846
OY 420 TCCTGGCTACACGAGCCCA-GAGATCAGGCGCTCGCATTTATATATGATGAGAGTATGATA 478
DB 847 TCCTGGCTACACGAGCCCA-GAGATCAGGCGCTCGCATTTATATATGATGAGAGTATGATA 906
OY 479 TGTCTTCT-CCTATGAGATGTGCTCTTA-CGAGTTGCTGTGAGACAGCGCC 527
DB 907 TGTCTTCTTCTTATGAGATGTGCTCTTACCGAGTGTGCTGACAGACAGCGCC 957

RESULT 8
BE899922 BE899922 562 bp mRNA linear EST 25-APR-2001
LOCUS 179602 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BE899922
ACCESSION BE899922
VERSION BE899922.1 GI:10387703
KEYWORDS EST.
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 562)
Smith,T.P.L., Grosse,M.W., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,M.W., Rohrer,G.A., Chitko-McKown,C.G.,
Peters,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013

TITLE

JOURNAL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mnscore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACACCTATGACCAT
BACKWARD: GTTTCACAGTACAGCAG
Plate: 82 row: F column: 5
Seq primer: ATTAGTGACACTATAG.

COMMENT

1. 562
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV Sport6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from day 20 and day 40 embryos."

BASE COUNT

131 a 156 c 175 g 100 t
ORIGIN

Query Match 21.0%; Score 459.6; DB 12; Length 562;
Best Local Similarity 88.6%; Pred. No. 7.9e-99;
Matches 498; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

OY 433 GCCCAGAGATGAGGCGCTCGCATTTATATATGATGAGAGTATGTTCTCTATGGA 492
DB 1 GCCCAGAGATGAGGCGCTCGCATTTATATATGATGAGAGTATGTTCTCTATGGA 60
OY 493 ATGTGCTCTACGAGTGTGCTGTACAGACGCGCTGCATGCGCCACACAGCTCCAG 552
DB 61 ATGTGCTGTACGAGTGTGCTGTACAGACGCGCGCTGCATGCGCCACACAGCTCCAG 120
OY 553 ATGGCAGAACCTGTCCAGAGGCGATCCGCGCGTCTGCGGCGACCGGAGGAAGTCCAG 612
DB 121 ATCTCCAGAACCTGTCCAGAGGCGATCCGCGCGTCTGCGGCGACCGGAGGAAGTCCAG 180
OY 613 TTCGCGGACTCAGAGCGCTCATGATGAGAGTCTGGACACTAAGCCAGAGAGGACCG 672
DB 181 TTCACACGACTCAGAGCGCTCATGATGAGAGTCTGGATACGAACCCGAGAGGACCA 240
OY 673 CTGCGCTGTGCGTGTGAGCCAGATGAGAGCCCGACTTTTGCACCTTCATGATGAA 732
DB 241 CTGCGCTGTGCGTGTGAGCCAGATGAGAGCCCGACTTTTGCACCTTCATGATGAA 300
OY 733 CTGTGCTGTGGGAGACAGACAGCTTCTTCATCCAGGCGCAGAGACAGTGTGTTG 792
DB 301 CTGTGCTGTGGGAGACAGACAGCTTCTTCATCCAGGCGCAGAGACAGTGTGTTG 360
OY 793 TTTTGGATGAGAAAGAGAGTCCAGAACTACAGGTGTGTGACAGAGAGAGGCGCTC 852
DB 361 TTTTGGATGAGAAAGAGAGTCCAGAACTACAGGTGTGTGACAGAGAGAGGCGCTC 420
OY 853 ATGAGGTGACAGAGATGTGCTCCCTGGGATGAAGGTAGTGTGACAGTCCAGTCCAG 912
DB 421 CTGGAAGTGCACACATGATGCTGTGCGGGAATGAAGTGTGACAGTCCAGTCCAG 480

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	8	GGATCAGCTACTGACTGCTCCATGCHCTCTACTCTCATGCCACCCCGCAGGCT	1066 bp	EST 12-JAN-2001				
QY	1696	GCCAGGTGCCCCCTCAAGCCTTCCCAAGCTCCCCAGCAATTTCTTCCAGTGTGCTTTC	1066 bp	EST 12-JAN-2001				
Db	68	GCCAGGTGCCCCCTCAAGCCTTCCCAAGCTCCCCAGCAATTTCTTCCAGTGTGCTTTC	1066 bp	EST 12-JAN-2001				
QY	1756	ACGACATGCGAGGAGCTGAGATGCTATCATATGAGCGCGGCTGCTCGACAGGCTGAG	1066 bp	EST 12-JAN-2001				
Db	128	ACGACATGCGAGGAGCTGAGATGCTATCATATGAGCGCGGCTGCTCGACAGGCTGAG	1066 bp	EST 12-JAN-2001				
QY	1816	CATGACCTGACCCCCCATGAGCGGAGAGCTTACAGCCAGCACTTCAGCGCGTGAAGATC	1066 bp	EST 12-JAN-2001				
Db	188	CATGACCTGACCCCCCATGAGCGGAGAGCTTACAGCCAGCACTTCAGCGCGTGAAGATC	1066 bp	EST 12-JAN-2001				
QY	1876	CTCGCCGCTGAGAGACTCATTTTGGGTGCCCAAGCGCGGTGAGATGTTATGCTATTGGC	1066 bp	EST 12-JAN-2001				
Db	248	CTCGCCGCTGAGAGACTCATTTTGGGTGCCCAAGCGCGGTGAGATGTTATGCTATTGGC	1066 bp	EST 12-JAN-2001				
QY	1936	CTGGAAGAAGATTCTGAAAGCCCAAGCGGCGGAGAGTCAATTCGCTTTAAAGCCGGAAG	1066 bp	EST 12-JAN-2001				
Db	308	CTGGAAGAAGATTCTGCGGCGCCAGCGGCGGAGAGTCAATTCGCTTTAAAGCCGGAAG	1066 bp	EST 12-JAN-2001				
QY	1996	CTGACTCCGATGCGGCTGCTGTGTGATGCTGCGCGGAGGCAAGAGACATGTTGTGTC	1066 bp	EST 12-JAN-2001				
Db	368	CTGACTCCGATGCGGCTGCTGTGTGATGCTGCGCGGAGGCAAGAGACATGTTGTGTC	1066 bp	EST 12-JAN-2001				
QY	2056	ACCTTTGAAATGAAACACACAGAGAGTGTGCTGCGCTCTGAGAGGCGGTGCGCCAGG	1066 bp	EST 12-JAN-2001				
Db	428	ACCTTTGAAATGAAACACACAGAGAGTGTGCTGCGCTCTGAGAGGCGGTGCGCCAGG	1066 bp	EST 12-JAN-2001				
QY	2116	GAGTTGCACTTTTCTACCAAGTCTTACGAGAGAGCTGGCGCGCTGAGAGCTTGCACTGC	1066 bp	EST 12-JAN-2001				
Db	488	GAGTTGCACTTTTCTACCAAGTCTTACGAGAGAGCTGGCGCGCTGAGAGCTTGCACTGC	1066 bp	EST 12-JAN-2001				
QY	2176	AAGAGAGG 2184	1066 bp	EST 12-JAN-2001				
Db	548	AAGAGAGG 556	1066 bp	EST 12-JAN-2001				

Query Match	24.8%	Score 541.6	DB 12	Length 1066
Best Local Similarity	92.1%	Pred. No. 2.9e-118		
Matches 604	Conservative 0	Mismatches 49	Indels 3	Gaps 3
BASE COUNT	263 a	323 c	283 g	197 t
ORIGIN				
Query Match				
1	ATGCTGAGGACACCTCGGGGCCACCGATGCCATGGAAGACTTCTCCGANTTCCGGCAGAG	60		
173	ATGCTGAGGACACCTCGGGGCCACCGATGCCATGGAAGACTTCTCCGANTTCCGGCAGAG	232		
61	GCGAGCATGCGACAGCGCTGAGACACCCGATCGGAGCGCTCATATGGCATCGATC	120		
233	GCGAGCATGCGACAGCGCTGAGACACCCGATCGGAGCGCTCATATGGCATCGATC	292		
121	CACCGC-CTGCTCTTGGCCTCGAGCTCGCGCGCGCTCAGACGCTCAACACCGTGTCTC	179		
293	CACCGACTGTGCTCTCGCGCTCGAGCTCGCGCGCGCTCAGACGCTCAACACCGTGTCTC	352		
180	CGAGACGCCGAGAGATTCTTCTCTTATATACCCCTGGGACATGCTCAACCAAAATATGC	239		
353	CGAGACGCCGAGAGATTCTTCTCTTATATACCCCTGGGACATGCTCAACCAAAATATGC	412		
240	CTACAGATGCGCTCGGGCGCTGAGCTGACATGACAAAGAAACATCATCTTGTGACCT	299		
413	CTACAGATGCGCTCGGGCGCTGAGCTGACATGACAAAGAAACATCATCTTGTGACCT	472		
300	GAAGTCGGAGCAACATTCGTGTGTGTGCTCCCTTGACGTCAAGAGCAGCATCAACATCAAGCT	359		
473	GAAGTCGGAGCAACATTCGTGTGTGTGCTCCCTTGACGTCAAGAGCAGCATCAACATCAAGCT	532		
360	ATTCGACTACGGGATTTGAGGCACTATTCATATAGGCGCCCTTAGGCGTGTGAGGGCAG	419		
533	ATTCGACTACGGGATTTGAGGCACTATTCATATAGGCGCCCTTAGGCGTGTGAGGGCAG	592		
420	TGCTGGCTACAGCGGCCAGAGATAGGCGCTGATTTATATGATGGAAGGTGATAT	479		
593	TGCTGGCTACAGCGGCCAGAGATAGGCGCTGATTTATATGATGGAAGGTGATAT	652		
480	GTTCTCTATGAAATGTGCTCTACAGAGTTGCTGACAGACAGCCCGCTGCACTGGGCCA	539		
653	TGTCTCTATGAAATGTGCTCTACAGAGTTGCTGACAGACAGCCCGCTGCACTGGGCCA	712		
540	CCACGAGCTCCAGATTGGCCAGAGAGCTGTCCAGAGGGCATCGCCCGGTTCTGGGGCAGCC	599		
713	CCACGAGCTCCAGATTGGCCAGAGAGCTGTCCAGAGGGCATCGCCCGGTTCTGGGGCAGCC	771		
600	GGAGGAGTGCAGTTCGGCGGCTCAGGCGCTCATATGATGAGAGTCTGGGAGACTA	655		
772	GGAGGAGTGCAGTTCGGCGGCTCAGGCGCTCATATGATGAGAGTCTGGGAGACTA	826		
RESULT 6				
LOCUS	BB613947	638 bp	mRNA	linear
DEFINITION	BB613947 RIKEN full-length enriched, 0 day neonate head Mus			
ACCESSION	musculus cDNA clone 483245J17 5', mRNA sequence.			
VERSION	BB613947			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hata,A.			

Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Jellord, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gunaratne, P., Yoon, V., Kowls, C., Martin, R., Lawrence, S., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM at: <http://Image.Illn.gov>
Series: IRK Plate: 19 Row: f Column: 19
This clone has the following problem: no 5' EST match.

FEATURES

source

1. 3509
/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="3463102"

/tissue_type="Cervix, carcinoma"

/clone_lib="NIH_MGC_12"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

BASE COUNT 810 a 994 c 953 g 752 t

ORIGIN

Query Match 87.7% Score 1915.4; DB 11; Length 3509;
Best Local Similarity 93.3% Pred. No. 0;

Matches 2074; Conservative 0; Mismatches 1; Indels 148; Gaps 1;

```

OY 110 GCATGACATCCACCGCTCTGCTTGGAGCTGCGCGCTCAGACCTCAACA 169
DB 1 GCATGACATCCACCGCTCTGCTTGGAGCTGCGCGCTCAGACCTCAACA 60
OY 170 CCGTGTGCGGAGAGCGCAGAGATCTCTTATACCCCTGGAGACATGCTACCC 229
DB 61 CCGTGTGCGGAGAGCGCAGAGATCTCTTATACCCCTGGAGACATGCTACCC 120
OY 230 AAAAAATAGCTACAGATCGCTCGGCGCTGCTACCTGCAAGAAAACATCATCT 289
DB 121 AAAAAATAGCTACAGATCGCTCGGCGCTGCTACCTGCAAGAAAACATCATCT 180
OY 290 TCTGTGACCTGAAGTCGGAACATCTTGTGTGTGCTTCTTACCTGCAAGAACATCA 349
DB 181 TCTGTGACCTGAAGTCGGAACATCTTGTGTGTGCTTCTTACCTGCAAGAACATCA 240
OY 350 ACATCAAGCTATCTACTACGAGGATTTGAGAGAGTATCCATGAGGCGCCAGAGCG 409
DB 241 ACATCAAGCTATCTACTACGAGGATTTGAGAGAGTATCCATGAGGCGCCAGAGCG 300
OY 410 TGGAGGCACTCTCTGCTACAGGCGCCAGAGATCGGCTGCAATTTATATGATAGA 469
DB 301 TGGAGGCACTCTCTGCTACAGGCGCCAGAGATCGGCTGCAATTTATATGATAGA 360
OY 470 AGGTGATATGTTCTCTTATGATATGATGATGATGATGATGATGATGATGATGAT 529
DB 361 AGGTGATATGTTCTCTTATGATATGATGATGATGATGATGATGATGATGATGAT 420
OY 530 CACTGGGCGACACAGGATCCAGATTGCAAGAGCTGCAAGGCGATCCCGCGGTTTC 589
DB 421 CACTGGGCGACACAGGATCCAGATTGCAAGAGCTGCAAGGCGATCCCGCGGTTTC 480
OY 590 TGGGCGACCGGAGAGATGCAAGTCCGCGAGCTGCAAGGCGCTCATGATGAGTCTGAG 649
DB 481 TGGGCGACCGGAGAGATGCAAGTCCGCGAGCTGCAAGGCGCTCATGATGAGTCTGAG 540
OY 650 ACACCTAACCCAGAGAGGAGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709
DB 541 ACACCTAACCCAGAGAGGAGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
OY 710 CTTTGGCCACCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 769
DB 601 CTTTGGCCACCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
OY 770 AGGGCCAGAGATACACCTGCTGTTTGGGATGAGAAAGAGAGATGAGAACTACAGG 829
DB 661 AGGGCCAGAGATACACCTGCTGTTTGGGATGAGAAAGAGAGATGAGAACTACAGG 720

```

```

OY 830 TGTGTAAACACAGAGAGGCGCTCATGAGAGTGCAGAGATGCTGCTGCTGCTGATGAAG 889
DB 721 TGTGTAAACACAGAGAGGCGCTCATGAGAGTGCAGAGATGCTGCTGCTGCTGATGAAG 780
OY 890 TGAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 949
DB 781 TGAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 840
OY 950 ACATATACACCTCAAGGCGATGCTGCTTAAACACACCCCAAGAGCGCTGATATC 1009
DB 841 ACATATACACCTCAAGGCGATGCTGCTTAAACACACCCCAAGAGCGCTGATATC 900
OY 1010 CAGCTGTGCTACCTGCTTCTTGGCGCTGCTTATTAAGAAATTCCTACCTGCTCT 1069
DB 901 CAGCTGTGCTACCTGCTTCTTGGCGCTGCTTATTAAGAAATTCCTACCTGCTCT 960
OY 1070 TAGCGGCGCTGCGCGATGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1129
DB 961 TAGCGGCGCTGCGCGATGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
OY 1130 ACAGTGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1189
DB 1021 ACAGTGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
OY 1190 AAGAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1249
DB 1081 AAGAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
OY 1250 AGGTGTGCTACAGCAATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309
DB 1141 AGGTGTGCTACAGCAATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
OY 1310 GCAGGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1369
DB 1201 GCAGGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
OY 1370 AGGCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1429
DB 1261 AGGCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
OY 1430 ACCACTCCACCACTACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1489
DB 1321 ACCACTCCACCACTACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
OY 1490 GGGAGATTTTCCGCTGCGGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549
DB 1381 GGGAGATTTTCCGCTGCGGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
OY 1550 ACCCAAAGGTGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1609
DB 1441 ACCCAAAGGTGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
OY 1610 TGGGAGCAGATCTGATCCACAGAGATCACTGACTGACTGCTGCTGCTGCTGCTGCT 1669
DB 1501 TGGGAGCAGATCTGATCCACAGAGATCACTGACTGACTGCTGCTGCTGCTGCTGCT 1560
OY 1670 ACTGCTCATCCCGACCGCGAGAGGCTGCGAGGCTGCGAGGCTGCGAGGCTGCGAGG 1729
DB 1561 ACTGCTCATCCCGACCGCGAGAGGCTGCGAGGCTGCGAGGCTGCGAGGCTGCGAGG 1620
OY 1730 CAAGTCTTCCAGTGTGCTTCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1789
DB 1621 CAAGTCTTCCAGTGTGCTTCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
OY 1790 CCGGCTGCTGCTGCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1849
DB 1681 CCGGCTGCTGCTGCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
OY 1850 GCGAGCACTGCGAGGCGTGAAGATCCGCGCGTCAAGAGAGAGAGAGAGAGAGAGAGAG 1909
DB 1741 GCGAGCACTGCGAGGCGTGAAGATCCGCGCGTCAAGAGAGAGAGAGAGAGAGAGAGAG 1800
OY 1910 GCGGTGAGATGATATGCTATGCTGCTGAGAGAGATTTGAGAGCCAGCGGCGGAG 1969

```


NIH-HGSC Project URL: <http://hgsc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-rt@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu

Oy 708 GACCTTTCACACCTTCATGATGACCTGCTGCTGGGAAGCAGACAGCCTTCTC-TCAT 766
|||||
Db 426 GACTTTCACACCTTCATGATGACCTGCTGCTGGGAAGCAGACAGCCTTCTCTCAT 485
|||||
Oy 767 CCCAGGCGCCAGAGTACACCGT-GGTGTTTGGATGAAA 806
|||||
Db 486 NCCAGGCGCCAGAGTACACTGTGGGTGTTTGGATGAAA 526
|||||

RESULT 12

AA572237
ID AA572237 standard; cDNA; 369 BP.

XX AC AA572237;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #8041.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KM food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR P-PSDB; AB608050.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.

XX PT Claim 1; SEQ ID No 8041; 103pp; English.

XX PS The invention relates to isolated polynucleotide (I) and
XX PS polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AA564197-AA594564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 369 BP; 88 A; 111 C; 94 G; 76 T; 0 other;

Query Match 4.7%; Score 103.4; DB 23; Length 369;
Best Local Similarity 99.0%; Pred. No. 6e-14;

Matches 104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 949 TACATCTACACCCCTAAGGCGATGCCCCCTAAACACACCCACAGCCTTGATCT 1008
|||||
Db 14 TACATCTACACCCCTAAGGCGATGCCCCCTAAACACACCCACAGCCTTGATCT 73
|||||
Oy 1009 CCAGCTGTCGTCACCTGCTTGGCCCGTCTGTTATTAATAAG 1053
|||||
Db 74 CCAGCTGTCGTCACCTGCTTGGCCCGTCTGTTATTAATAAG 118
|||||

RESULT 13

AA587109
ID AA587109 standard; cDNA; 1625 BP.

XX AC AA587109;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #22913.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KM food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR P-PSDB; AB622922.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.

XX PT Claim 1; SEQ ID No 22913; 103pp; English.

XX PS The invention relates to isolated polynucleotide (I) and
XX PS polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AA564197-AA594564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1625 BP; 330 A; 491 C; 471 G; 333 T; 0 other;

PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 25-SEP-2000; 2000US-0234999.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235835.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241121.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249269.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250161.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and
 metastasis

PT Disclosure; SEQ ID NO 38919; 3071bp + Sequence Listing; English.

XX AA654951 to AA64702 encode the human immune/hematopoietic antigen (I)
 amino acid sequences given in AA682170 to AA691921. (I) have cytosolic
 activity, and can be used in gene therapy and vaccine production. (I)
 proteins and polynucleotides may be used in the prevention, diagnosis and
 treatment of diseases associated with inappropriate (I) expression. For
 example, they may be used to treat disorders associated with decreased
 expression by rectifying mutations or deletions in a patient's genome
 that affect the activity of (I) by expressing inactive proteins or to
 supplement the patients own production of (I). Additionally, (I)
 polynucleotides may be used to produce the secreted (I), by inserting
 the nucleic acids into a host cell and culturing the cell to express the
 protein. (I) proteins and polynucleotides may be used to prevent,
 diagnose and treat immune/hematopoietic-related diseases, especially
 cancers and cancer metastases of haematopoietic-derived cells. AA64703
 to AA67694 represent human immune/hematopoietic antigen genomic
 sequences from the present invention. AA654942 to AA654950 and AA682169
 represent sequences used in the exemplification of the present invention.

XX Sequence 18287 BP; 4371 A; 4858 C; 4942 G; 4116 T; 0 other;

Query Match 39.3%; Score 858.8; DB 22; Length 18287;

Best Local Similarity 96.0%; Pred. No. 3,2e-185; Mismatches 37; Indels 0; Gaps 0;

Matches 881; Conservative 0;

QY 1052 AGAATTCCTACCTGCTCTAGGCGGCTGCGCATGCGCTTGTGCTGTTCCCGTGG 1111
 DB 13525 AGAATTCCTACCTGCTCTAGGCGGCTGCGCATGCGCTTGTGCTGTTCCCGTGG 13584
 QY 1112 TCGGGGGACCCCAAGAGACAGCTGCTCTACCTGCTGTCACACAGCCAAAGGTCCA 1171
 DB 13585 TCGGGGGACCCCAAGAGACAGCTGCTCTACCTGCTGTCACACAGCCAAAGGTCCA 13644
 QY 1172 AGTTGAGCATCGCGATGAAGACGCGGACAGAACCCCTAACCCAGGAGGCGATGGAGG 1231
 DB 13645 AGTTGAGCATCGCGATGAAGACGCGGACAGAACCCCTAACCCAGGAGGCGATGGAGG 13704
 QY 1232 TGGTCAACAGCGGCTCTGAGGCTGTGATACGCAATGGCGGCGCTCTGTATCGACT 1291
 DB 13705 TGGTCAACAGCGGCTCTGAGGCTGTGATACGCAATGGCGGCGCTCTGTATCGACT 13764
 QY 1292 GTGCTCCCTGGAGATCTCAGGCGGCTGAGGCTTACATGGCCCCCTCATGTTACT 1351

CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent, ex-
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK44703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 XX Sequence 18287 BP; 4371 A; 4858 C; 4942 G; 4116 T; 0 other;

Query Match 39.38; Score 858.8; DB 22; Length 18287;
 Best Local Similarity 96.0%; Pred. No. 3.2e-185;
 Matches 881; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1052 AGAATTCCTACCTGCTCTAGAGGGGCTCGCGAGGGCTTGCTGCTGTTCCCGG 1111
 DB 13555 AGAATTCCTACCTGCTCTAGAGGGGCTCGCGAGGGCTTGCTGCTGTTCCCGG 13584
 QY 1112 TCGGGGACCCCAAGAGACAGCTGCTCTACCTGCTGTCACACAGCAAGAGTCCA 1171
 DB 13585 TCGGGGACCCCAAGAGACAGCTGCTCTACCTGCTGTCACACAGCAAGAGTCCA 13644
 QY 1172 AGTTAGCATGCGGATGAGACGACGCGAGACCCCTACCACTGGAAGCCATGAGG 1231
 DB 13645 AGTTAGCATGCGGATGAGACGACGCGAGACCCCTACCACTGGAAGCCATGAGG 13704
 QY 1232 TGGTCAACAGGGGCTCTAGAGGCTGTGATAGCAATGGGCGGGGCTCTGTCATGAG 1291
 DB 13705 TGGTCAACAGGGGCTCTAGAGGCTGTGATAGCAATGGGCGGGGCTCTGTCATGAG 13764
 QY 1292 GTGCTCTCCCTGAGATCTGACAGGCGGCTGAGACCTTACATGGGCCCTCCATGTTAGT 1351
 DB 13765 GTGCTCTCCCTGAGATCTGACAGGCGGCTGAGACCTTACATGGGCCCTCCATGTTAGT 13824
 QY 1352 CAGTGTGTGACCTCTGAGGGCAAGGGAGAGAGTGTCTGTGCTGTGATGACAAG 1411
 DB 13825 CAGTGTGTGACCTCTGAGGGCAAGGGAGAGAGTGTCTGTGCTGTGATGACAAG 13884
 QY 1412 CCAACTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1471
 DB 13885 CCAACTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13944
 QY 1472 GGGTCCCAAGGCGCTCTAGAGGACATGTTCCGCTGCGGCTTGGACAGCAAGCCCGG 1531
 DB 13945 GGGTCCCAAGGCGCTCTAGAGGACATGTTCCGCTGCGGCTTGGACAGCAAGCCCGG 14004
 QY 1532 CAGCCAGCCACAGGCGCAACCCAAAGGTGCTGAGGGGAGTCCATCGCGGAGCTGAGCA 1591
 DB 14005 CAGCCAGCCACAGGCGCAACCCAAAGGTGCTGAGGGGAGTCCATCGCGGAGCTGAGCA 14064
 QY 1592 TCATGTAGAGTAGAGAGTGGGAGAGATGCTGATCCACAGAGATGATGATGATGATGAT 1651
 DB 14065 TCATGTAGAGTAGAGAGTGGGAGAGATGCTGATCCACAGAGATGATGATGATGATGAT 14124
 QY 1652 ACTGCTCATGCTCTCTACTCTCTATCCCAAGCCGCAAGGCTGCAAGTCCCTCA 1711
 DB 14125 ACTGCTCATGCTCTCTACTCTCTATCCCAAGCCGCAAGGCTGCAAGTCCCTCA 14184
 QY 1712 GCTTCCCAAGCTCCCAAGTCTTCCAGTGTGCTTTCACAGCACTGCGAGGACT 1771
 DB 14185 GCTTCCCAAGCTCCCAAGTCTTCCAGTGTGCTTTCACAGCACTGCGAGGACT 14244
 QY 1772 CAGACATGCTACATACGCGCGGCTGCTCGAGAGGTCTAGATGATGATGATGATGATGAT 1831
 DB 14245 CAGACATGCTACATACGCGCGGCTGCTCGAGAGGTCTAGATGATGATGATGATGATGAT 14304
 QY 1832 TGGAGGGGAGACTTTCAGCAGCACTGACAGGCGGTGAAGATCTCGCGGTGAGAGACC 1891
 DB 14305 TGGAGGGGAGACTTTCAGCAGCACTGACAGGCGGTGAAGATCTCGCGGTGAGAGACC 14364
 QY 1892 TCATTGGGTCCCAAGGCGGCTGAGATGTTATGCTATTGCGCTGAGAGAGATTTCTG 1951
 DB 14365 TCATTGGGTCCCAAGGCGGCTGAGATGTTATGCTATTGCGCTGAGAGAGATTTCTG 14424

DB 14365 TCATTGGGTCCCAAGGCGGCTGAGATGTTATGCTATTGCGCTGAGAGAGATTTCTG 14424
 QY 1952 AAGCCAGCGGGGCGGAG 1969
 DB 14425 CTCCTGCTCTGGGGACAG 14442
 RESULT 10
 ID AAK84107
 AC AAK84107 standard; DNA: 18287 BP.
 AC AAK84107;
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38919.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytosolic; gene therapy; vaccine; metastasis; ds.
 OS Homo sapiens.
 PN W0200157182-A2.
 PD
 XX
 PD
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0230963.
 PR 26-JUL-2000; 2000US-0230964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.

PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0255678.
 (HUMA-) HDMA GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 DR WPI; 2001-483426/52.
 XX
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure: SEQ ID NO 38916; 3071pp + Sequence Listing; English.
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK67694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 XX
 SO Sequence 18286 BP; 4372 A; 4856 C; 4941 G; 4117 T; 0 other;
 Query Match 39.3%; Score 858.8; DB 22; Length 18286;
 Best Local Similarity 96.0%; Pred. No. 3.2e-185;
 Matches 881; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

DB 13884 CCAACTCCTGGTGTATGATCACTCCACCTACACAGCTGTGTGCGGATCTGTGG 13943
 QY 1472 GGGTCCCAACGCCCTCCAGGAGCATGTTCCCGTGGCCCTTGACACAGCAACCCCGG 1531
 DB 13944 GGGTCCCAACGCCCTCCAGGAGCATGTTCCCGTGGCCCTTGACACAGCAACCCCGG 14003
 QY 1532 CAGCCAGCCACAGGAGCAACCCCAAGAGTCTGAGGAGGAGTCCATCCGAGAGTGA 1591
 DB 14004 CAGCCAGCCACAGGAGCAACCCCAAGAGTCTGAGGAGGAGTCCATCCGAGAGTGA 14063
 QY 1592 TCATGTACAGTGAAGAGTGGGACGACAGATCCTGATCCACAGCAATCACTACTACT 1651
 DB 14064 TCATGTACAGTGAAGAGTGGGACGACAGATCCTGATCCACAGCAATCACTACTACT 14123
 QY 1652 ACTGCTCCATGTCT 1711
 DB 14124 ACTGCTCCATGTCT 14183
 QY 1712 GCCTCCCAAGTCCCAAGCAATGCTTCCAGTGTGCTTCCACAGCAATCACTACTACT 1771
 DB 14184 GCCTCCCAAGTCCCAAGCAATGCTTCCAGTGTGCTTCCACAGCAATCACTACTACT 14243
 QY 1772 CAGACATGCTACATACGCGCGGTGCTGCTCCGACAGTGTGACATGACCTGACCCCA 1831
 DB 14244 CAGACATGCTACATACGCGCGGTGCTGCTCCGACAGTGTGACATGACCTGACCCCA 14303
 QY 1832 TGGAGGGGAGACCTTTCACGACGACCTGACAGCGCTGAAGATCTCGCGCTGACAGACC 1891
 DB 14304 TGGAGGGGAGACCTTTCACGACGACCTGACAGCGCTGAAGATCTCGCGCTGACAGACC 14363
 QY 1892 TCATTTGGTGGTCCCAAGGAGCGGTGAGATGATGATGATGATGATGATGATGATGAT 1951
 DB 14364 TCATTTGGTGGTCCCAAGGAGCGGTGAGATGATGATGATGATGATGATGATGATGAT 14423
 QY 1952 AAGCCAGCGGGGCGAG 1969
 DB 14424 CTCCTGCTCTGGGACAG 14441
 RESULT 9
 ID AAK70892
 AC AAK70892 standard; DNA; 18287 BP.
 AC AAK70892;
 DT 06-NOV-2001 (first entry)
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25704.
 DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW Cytostatic; gene therapy; vaccine; metastasis; ds.
 OS Homo sapiens.
 PN WO200157182-A2.
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0188874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.

```
RESULT 4
AA684104
ID AA684104 standard; DNA; 18286 BP.
XX
AC AAR64104
XX
UT 07-NOV-2001 (first entry)
XX
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:38916.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KV Cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
```

```
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249247.
PR 17-NOV-2000; 2000US-0249248.
PR 17-NOV-2000; 2000US-0249249.
PR 17-NOV-2000; 2000US-0249250.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
```

PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0246207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251899.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259478.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI: 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure: SEQ ID NO 25701; 3071bp + Sequence listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169

CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 18286 BP; 4372 A; 4856 C; 4941 G; 4117 T; 0 other;
 Query Match 39.3%; Score 858.8; DB 22; Length 18286;
 Best Local Similarity 96.0%; Pred. No. 3,2e+185;
 Matches 881; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
 OY 1052 AGAATTCCTACTGCTCTTACGCGGCTCCGATGGGCTTGGCTGTGTTCCCTGG 1111
 DB 13524 AGAATTCCTACTGCTCTTACGCGGCTCCGATGGGCTTGGCTGTGTTCCCTGG 13583
 OY 1112 TGGGGGACCCCAAGGACAGCTGCTCTTACCTGTGTACACACAGCCACAGTCCA 1171
 DB 13584 TGGGGGACCCCAAGGACAGCTGCTCTTACCTGTGTACACACAGCCACAGTCCA 13643
 OY 1172 AGTTCAGATCGCGGATGAAGACGCGGACAGAACCCCTACCACTGAAGGCGCATGAGG 1231
 DB 13644 AGTTCAGATCGCGGATGAAGACGCGGACAGAACCCCTACCACTGAAGGCGCATGAGG 13703
 OY 1232 TGGTCAACAGCGGCTCTGAGGTCTGTACAGCATGGCCGGGCTCTGTATCATGACT 1291
 DB 13704 TGGTCAACAGCGGCTCTGAGGTCTGTACAGCATGGCCGGGCTCTGTATCATGACT 13763
 OY 1292 GTGCTCCCTGGAGATCTGACGGGGGCTGGAGCCCTACATGGCCCTCCATGGTTAGT 1351
 DB 13764 GTGCTCCCTGGAGATCTGACGGGGGCTGGAGCCCTACATGGCCCTCCATGGTTAGT 13823
 OY 1352 CAGTCGTGACAGCTCTGAGGGGACAGAGGAGAGTGTGTGCTGTGATGACAAAG 1411
 DB 13824 CAGTCGTGACAGCTCTGAGGGGACAGAGGAGAGTGTGTGCTGTGATGACAAAG 13883
 OY 1412 CCAACTCTTGGTGTATGTACACATCCACCTACCACTGTGTGCGCGGACTTCTGCG 1471
 DB 13884 CCAACTCTTGGTGTATGTACACATCCACCTACCACTGTGTGCGCGGACTTCTGCG 13943
 OY 1472 GGGTCCCGACGCCCTCTGAGGGGACATGTTCCCGTGGGCGCTTGGACACGGAACCCCGG 1531
 DB 13944 GGGTCCCGACGCCCTCTGAGGGGACATGTTCCCGTGGGCGCTTGGACACGGAACCCCGG 14003
 OY 1532 CAGCAGCCACAGCGGCCAACCCAAAGGTGCTGAGGGGAGTCCATCCGAGACGTGACA 1591
 DB 14004 CAGCAGCCACAGCGGCCAACCCAAAGGTGCTGAGGGGAGTCCATCCGAGACGTGACA 14063
 OY 1592 TCATGTACAGTGAAGAGTGGGACAGCATCTGTATCCACAGGAATCACTCATGACT 1651
 DB 14064 TCATGTACAGTGAAGAGTGGGACAGCATCTGTATCCACAGGAATCACTCATGACT 14123
 OY 1652 ACTGCTCATGCTCTCTACACGCTCATACCCCAACCCCGCAGGCTGCCAGGTCCCTCAA 1711
 DB 14124 ACTGCTCATGCTCTCTACACGCTCATACCCCAACCCCGCAGGCTGCCAGGTCCCTCAA 14183
 OY 1712 GCTTCCCGAGCTCCCAAGCAAGTCTTCCAGATGAGTCTTCCACGAGCTGCGAGACT 1771
 DB 14184 GCTTCCCGAGCTCCCAAGCAAGTCTTCCAGATGAGTCTTCCACGAGCTGCGAGACT 14243
 OY 1772 CAGACATCTACATACGCCCGCGTCTCTCCAGACAGTCTGACATGACCTGACCCCA 1831
 DB 14244 CAGACATCTACATACGCCCGCGTCTCTCCAGACAGTCTGACATGACCTGACCCCA 14303
 OY 1832 TGGACGGGGGAGACCTTACAGCAGCACTGCGAGCGGCGGAGATCTCTGCTGAGAGCC 1891
 DB 14304 TGGACGGGGGAGACCTTACAGCAGCACTGCGAGCGGCGGAGATCTCTGCTGAGAGCC 14376
 OY 1892 TCATTTGGGTCCCGACAGCGCGGTGAGATGTTATCTGATGAGCTGGAGAGATCTG 1951
 DB 14364 TCATTTGGGTCCCGACAGCGCGGTGAGATGTTATCTGATGAGCTGGAGAGATCTG 14423
 OY 1952 AAGCCCAAGCGGCGCGAG 1969
 DB 14424 CTCCTGCTGTGGGACAG 14441

QY	1741	AGTGTGCTTTTCCACACGACATCGGAGACTAGACATGCTACATACGCCGGGTGTGCC	1800
Db	5825	AGGTGTGCTTTTCCACACGACTCGCGAGACTAGACATGCTACATACGCCGGGTGTGCC	5884
QY	1801	TCCGACAGGTCTAGATGATGACGTGACCCCCATGAGCGGGAGACCTTAGCCAGACTCG	1860
Db	5885	TCGACAGGTCTAGCATGACCTGACCCCCATGAGCGGGAGAACCTTAGCCAGACTCG	5944
QY	1861	CAGGCCGTGAAGATCCTCGCCGTCAGAGACTTCATTTTGGGTCCCCAGGCGGGTGGAGAT	1920
Db	5945	CAGGCCGTGAAGATCCTCGCCGTCAGAGACTTCATTTTGGGTCCCCAGGCGGGTGGAGAT	6004
QY	1921	GTTATGCTCATTTGGCGTGGAGAGAAATTCGAAACCCAGCGGGGCCGATCTTGGCGTC	1980
Db	6005	GTTATGCTCATTTGGCGTGGAGAGAAATTCGCGGCCACGGGGGCCGATCTTGGCGTC	6064
QY	1981	TTAAAGCCCCGAGAGCTGACTCCGCATGGGGTGTGTTGGATGTCTCCGTGGTGGCAAG	2040
Db	6065	TTAAAGCCCCGAGAGCTGACTCCGCATGGGGTGTGTTGGATGTCTCCGTGGTGGCAAG	6124
QY	2041	GACACTGTGTGTGCTGACTTTGAAATGAAACACAGAGATGTGTGCTGGCCGTCTGGAG	2100
Db	6125	GACACTGTGTGTGCTGACTTTGAAATGAAACACAGAGATGTGTGCTGGCCGTCTGGAG	6184
QY	2101	GGCTGGGGCGCCGAGAGATTGCAATTTTCTACAGTCTCTACGAGAGCTGGGCGGCTG	2160
Db	6185	GGCTGGGGCGCCGAGAGATTGCAATTTTCTACAGTCTCTACGAGAGCTGGGCGGCTG	6244
QY	2161	GAGGCTTGCACTCGCAGAGAAAG	2184
Db	6245	GAGGCTTGCACTCGCAGAGAAAG	6268

XX Claim 5; Page 192-194; 210bp; English.
PS
XX The invention relates human kinases (PKIN) and their corresponding
CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
CC treating and preventing cancer, an immune system disorder (e.g.,
CC acquired immune deficiency syndrome (AIDS), Addison's disease, allergy,
CC asthma, atherosclerosis, multiple sclerosis, psoriasis), disorders
CC affecting growth and development (e.g., arteriosclerosis, cirrhosis,
CC hepatitis), cardiovascular disorder (e.g., hypertension, myocardial
CC infarction, Goodpasture's syndrome), and a lipid disorder (e.g., fatty
CC liver, Gaucher's disease, Niemann-Pick's disease, hypercholesterolemia,
CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
CC condition or a disease associated with the expression of PKIN in a
CC biological sample. A composition comprising PKIN or an agonist or
CC antagonist of PKIN is useful for treating a disease or conditional PKIN.
CC PKIN is useful in a number of drug screening techniques and to analyse
CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
CC knock-in humanised animals or transgenic animals to model human diseases,
CC and in somatic or germline gene therapy. The present sequence is human
CC PKIN cDNA.
XX
SQ Sequence 7093 BP; 1554 A; 2102 C; 2027 G; 1410 T; 0 other;
Query Match 88.7%; Score 1936.8; DB 24; Length 7093;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1;
QY 1 ATGCTGAGGACATGCGGGCCACGATGAGAGAACTTCTCCAGTTCGGGAGAG 60
DB 4202 ATGCTGAGGACATGCGGGCCACGATGAGAGAACTTCTCCAGTTCGGGAGAG 4261
QY 61 GCCAGATGCTGACAGCGCTGACAGACCCCTGATGCTGCGGCTATGCGATCAGCATC 120
DB 4262 GCCAGATGCTGACAGCGCTGACAGACCCCTGATGCTGCGGCTATGCGATCAGCATC 4321
QY 121 CACCGCTCTGCTGCGCTGAGAGCTGCGCGCTCAGCAGCGCTCAACACCGTGTGCC 180
DB 4322 CACCGCTCTGCTGCGCTGAGAGCTGCGCGCTCAGCAGCGCTCAACACCGTGTGCC 4381
QY 181 GAGAAAGCCAGAGATTTCTTCTTATACCCCTGGAGACATGCTCACCAGAAATATGCC 240
DB 4382 GAGAAAGCCAGAGATTTCTTCTTATACCCCTGGAGACATGCTCACCAGAAATATGCC 4441
QY 241 TACGAGATGCGCTGCGGCTGCGCTGACGACAGAGAAATATATCTTCTGTAAGCTG 300
DB 4442 TACGAGATGCGCTGCGGCTGCGCTGACGACAGAGAAATATATCTTCTGTAAGCTG 4501
QY 301 AAGTGGACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 4502 AAGTGGACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4561
QY 361 TCTGACTACGGGATTTGAGGAGTCAATTCATGAGGGGCGCTTAGCGGTGAGGGCACT 420
DB 4562 TCTGACTACGGGATTTGAGGAGTCAATTCATGAGGGGCGCTTAGCGGTGAGGGCACT 4621
QY 421 CCTGCTACGAGGCGCCAGAGATGAGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 4622 CCTGCTACGAGGCGCCAGAGATGAGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4681
QY 481 TTTCTCTATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 4682 TTTCTCTATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4741
QY 541 CACGAGCTCAGATTTGCAAGAGCTGTCCAAAGGCGATCCGCGGTTCTGCGGAGCGG 600
DB 4742 CACGAGCTCAGATTTGCAAGAGCTGTCCAAAGGCGATCCGCGGTTCTGCGGAGCGG 4801
QY 601 GAGGAGAGGAGCTGCGGAGCTGAGAGGCGCTGATGAGAGTGTGCTGGAGACTAAGCA 660
DB 4802 GAGGAGAGGAGCTGCGGAGCTGAGAGGCGCTGATGAGAGTGTGCTGGAGACTAAGCA 4861

QY 661 GAGAGAGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 4862 GAGAGAGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4921
QY 721 TTTCTATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 4922 TTTCTATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4981
QY 781 TACACCGTGTGCTTGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 4982 TACACCGTGTGCTTGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5041
QY 841 GAGAGAGCGGCTCAG 900
DB 5042 GAGAGAGCGGCTCAG 5101
QY 901 CTCAGGCTCAGAGATCCCTGTGAGACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 5102 CTCAGGCTCAGAGATCCCTGTGAGACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5135
QY 961 CTCAGGAGATGTGCGCTTAAACACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 5136 ----- 5135
QY 1021 ACCTGCTTCTTGCGCGGCTGCTGCTTAAAGAGATTCCTACCTGCTTACGCGGCTC 1080
DB 5136 ----- AGAATTCCTACCTGCTTACGCGGCTC 5184
QY 1081 GCCAGTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 5185 GCCAGTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5224
QY 1141 TACCTGCTCAGACACAGCCAG 1200
DB 5225 TACCTGCTCAGACACAGCCAG 5284
QY 1201 CAGAAACCCCTACAG 1260
DB 5285 CAGAAACCCCTACAG 5344
QY 1261 AGCAATGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 5345 AGCAATGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5404
QY 1321 GAGCCCTACATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 5405 GAGCCCTACATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5464
QY 1381 GAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 5465 GAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5524
QY 1441 ACCTACACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 5525 ACCTACACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5584
QY 1501 CCGGTGCGGCGCTTGTGAGACAG 1560
DB 5585 CCGGTGCGGCGCTTGTGAGACAG 5644
QY 1561 CCGTGAAGGAGATCCATGCGAG 1620
DB 5645 CCGTGAAGGAGATCCATGCGAG 5704
QY 1621 ATCTGATGCTACAGAGATCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1680
DB 5705 ATCTGATGCTACAGAGATCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 5764
QY 1681 CCACCGGCGAGAGGCTGCGAGGCTGCGAGGCTGCGAGGCTGCGAGGCTGCGAGGCTGCGAGGCT 1740
DB 5765 CCACCGGCGAGAGGCTGCGAGGCTGCGAGGCTGCGAGGCTGCGAGGCTGCGAGGCTGCGAGG 5824

Db 4816 GAGAAAGGCGCTCATGAGAGTGCAGAGATGTCTGCCCTGGATGAGTGAAGTGCACG 4875
 QY 901 CTCAGGTCACAGAGATCCCTGTGTGACACCGACCGAGACAGAAAAATCTACATCTACACC 960
 Db 4876 CTCAGGTCACAGAGATCCCTGTGTGACACCGACCGACCGACCGACCGACCGACCG 4909
 QY 961 CTCAGGTCACAGAGATCCCTGTGTGACACCGACCGACCGACCGACCGACCGACCG 1020
 Db 4910 ----- 4909
 QY 1021 ACCCTCTTCTGGCGCTGTCTATTAAGAATTCCTACCTGTCTTACGCGGCTC 1080
 Db 4910 ----- 4938
 QY 1081 GCGGATGGGCTGT 1140
 Db 4939 GCGGATGGGCTGT 4998
 QY 1141 TACCTGT 1200
 Db 4999 TACCTGT 5058
 QY 1201 CAGAACCCCTACACAGTGAAGGCAATGTGATGATGATGATGATGATGATGATGATGAT 1260
 Db 5059 CAGAACCCCTACACAGTGAAGGCAATGTGATGATGATGATGATGATGATGATGATGAT 5118
 QY 1261 AGCAATGGGCGGCGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
 Db 5119 AGCAATGGGCGGCGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5178
 QY 1321 GAGCCCTACATGGGCGGCGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
 Db 5179 GAGCCCTACATGGGCGGCGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5238
 QY 1381 GAGGAGGT 1440
 Db 5239 GAGGAGGT 5298
 QY 1441 ACCTACCACTGT 1500
 Db 5299 ACCTACCACTGT 5358
 QY 1501 CCCGT 1560
 Db 5359 CCCGT 5418
 QY 1561 CCTGAGGAGGAGT 1620
 Db 5419 CCTGAGGAGGAGT 5478
 QY 1621 ATCTGTATCCACAGGAGT 1680
 Db 5479 ATCTGTATCCACAGGAGT 5538
 QY 1681 CCACCCCGGAGGAGT 1740
 Db 5539 CCACCCCGGAGGAGT 5598
 QY 1741 AGT 1800
 Db 5599 AGT 5658
 QY 1801 TCCGAGGAGT 1860
 Db 5659 TCCGAGGAGT 5718
 QY 1861 CAGGCGGTGAAGATCTCTGCGGTGAGAGATCTCTGCGGTGAGAGATCTCTGCGGTGAGAG 1920
 Db 5719 CAGGCGGTGAAGATCTCTGCGGTGAGAGATCTCTGCGGTGAGAGATCTCTGCGGTGAGAG 5778
 QY 1921 GTTATGTATGT 1980
 Db 5779 GTTATGTATGT 5838

QY 1981 TTAAGAGCCGAGAGAGT 2040
 Db 5839 TTAAGAGCCGAGAGAGT 5898
 QY 2041 GACACTGT 2100
 Db 5899 GACACTGT 5958
 QY 2101 GCGT 2160
 Db 5959 GCGT 6018
 QY 2161 GAGCGT 2184
 Db 6019 GAGCGT 6042
 Db 6019 GAGCGT 6042
 RESULT 6
 AAD38851
 ID AAD38851 standard; cDNA; 7093 BP.
 XX
 AC AAD38851;
 XX
 DT 23-SEP-2002 (first entry)
 XX
 DE Human kinase (PKIN)-8 cDNA.
 XX
 KW Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
 KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
 KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
 KW development; hepatitis; cardiovascular; hypertension; drug screening;
 KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
 KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
 KW hypercholesterolemia; obesity; gene therapy; cytostatic; anti-HIV;
 KW neuroprotective; hepatotropic; hypotensive; cardiant; nephrotropic;
 KW hyperlipidemia; enzyme; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FT 227..6271
 FT CDS /tag="a
 FT /product="Human kinase (PKIN)-8"
 FT
 PN WO200233099-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 20-OCT-2001; 2001WO-US47728.
 XX
 PR 20-OCT-2000; 2000US-242410P.
 PR 27-OCT-2000; 2000US-244068P.
 PR 03-NOV-2000; 2000US-245708P.
 PR 09-NOV-2000; 2000US-247672P.
 PR 16-NOV-2000; 2000US-249565P.
 PR 22-NOV-2000; 2000US-252730P.
 PR 01-DEC-2000; 2000US-250807P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Gururajan R, Baughn MR, Walla NK, Elliott VS, Xu Y, Arvizu C;
 PI Yao MG, Ramkumar J, Ding L, Tang YT, Hatalla AD, Nguyen DB;
 PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM;
 PI Lal PG, Recipon SA, Lu DM, Borowsky ML, Thornton M, Swarnaker A;
 PI Thangavelu K, Khan FA, Ison CH;
 XX
 DR WPI: 2002-454603/48.
 DR P-PSDB; AAE24137.
 XX
 PT New human kinase polypeptide, for diagnosing, preventing and treating
 PT cancer, immune system disorders, growth and development disorders,
 PT cardiovascular disorders and lipid disorders

QY 1979 TCTTAAAGCCGAGAGTACTCCGATGGGTGCTGTGGATGCTGCGGTGGCAA 2038
 CC |||||||
 DB 2425 CCTTAAAGCCGAGAGTACTCCGATGGGTGCTGTGGATGCTGCGGTGGCAA 2484
 CC |||||||
 QY 2039 AGGACACTGTGTGTGC-ACCTTGAATAAAGAACAGAGTGTGCTGGCCGTCTG 2097
 CC |||||||
 DB 2485 AGGACACTGTGTGTGCACCTTTGAAAAAGAACAGAGTGTGCTGGCCGTCTG 2544
 CC |||||||
 QY 2098 AGGGCTGGGGCCGAGAGTTCGACATTTCTACACATCTCTACAGAGTGGCCGG 2157
 CC |||||||
 DB 2545 AGGGCTGGGGCCGAGAGTTCGACATTTCTACACATCTCTACAGAGTGGCCGG 2604
 CC |||||||
 QY 2158 CTGAGAGCTTGACATCGCAGAGAAAG 2183
 CC |||||||
 DB 2605 CTGAGAGCTTGACATCGCAGAGAAAG 2630
 CC |||||||
 RESULT 5
 AAH77994
 ID AAH77994 standard; DNA; 6045 BP.
 AC AAH77994;
 XX
 DT 13-NOV-2001 (first entry)
 XX
 DE Nucleotide sequence of human protein kinase SGK258.
 XX
 KM Human, protein kinase; cancer; immune disease; cardiovascular disease;
 KM brain disease; neuronal disease; Alzheimer's disease; chromosome 15;
 KM Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KM peripheral nervous system disease; migraine; pain; sexual dysfunction;
 KM infection; ocular disease; migraine; pain; sexual dysfunction;
 KM mood disorder; attention disorder; cognition disorder; hypotension;
 KM hypertension; psychotic disorder; dyskinesia; transplant rejection; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..6045
 FT /tag- a
 FT /product= "protein kinase"
 FT
 PN WO20016594-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 02-MAR-2001; 2001WO-US06838.
 XX
 PR 06-MAR-2000; 2000US-0187150.
 PR 29-MAR-2000; 2000US-0193404.
 PR 13-NOV-2000; 2000US-0247013.
 XX
 PA (SUGEN) SUGEN INC.
 XX
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 DR WPI: 2001-536777/59.
 DR P-PSDB; AAG67395.
 XX
 PT Nucleic acids capable of encoding human polypeptides having a kinase or
 PT kinase-like activity, useful for diagnosing a disease selected from
 PT cancers, cardiovascular disease and neuronal-associated diseases (e.g.
 PT Alzheimer's disease)
 XX
 PS Example 1; Fig 1E-F; 201pp; English.
 XX
 CC The present sequence encodes a human protein kinase. The gene is
 CC located on chromosome 15. The kinase polypeptides are useful for
 CC diagnosing a disease or disorder selected from cancers (e.g. cancers
 CC of tissues and cancers of hematopoietic origin), immune-related diseases
 CC and disorders, cardiovascular disease, brain or neuronal-associated
 CC diseases (e.g. Alzheimer's disease, Parkinson's disease, multiple

CC sclerosis); metabolic disorders, peripheral nervous system diseases,
 CC amyotrophic lateral sclerosis, viral infections, infections caused by
 CC prions, infections caused by bacteria, infections caused by fungi,
 CC ocular diseases, migraines, pain, sexual dysfunction, mood disorders,
 CC attention disorders, cognition disorders, hypotension, hypertension,
 CC psychotic disorders, dyskinesias, and organ transplant rejection.
 CC Kinase inhibitors are useful for treating diseases and disorders
 CC described above.
 XX
 SQ Sequence 6045 BP; 1359 A; 1779 C; 1707 G; 1200 T; 0 other;
 Query Match 88.7%; Score 1936.8; DB 22; Length 6045;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 2065; Conservative 0; Mismatches 2; Indels 117; Gaps 1;
 QY 1 ATGCTGAGGACCTGCGGGCCAGCAGATGCGATGAAGAACTTCCGAGTCCGGAGAG 60
 DB 3976 ATGCTGAGGACCTGCGGGCCAGCAGATGCGATGAAGAACTTCCGAGTCCGGAGAG 4035
 QY 61 GCCAGATGCTGACAGCGGTGTGAGACACCCCTGATCTGTGGCTATCGGATCAGATC 120
 DB 4036 GCCAGATGCTGACAGCGGTGTGAGACACCCCTGATCTGTGGCTATCGGATCAGATC 4095
 QY 121 CACCCGCTGCTGCTGCGCTGTGAGAGTGGCGCGCTCAGAGGCTCAACACGCTGTGTC 180
 DB 4096 CACCCGCTGCTGCTGCGCTGTGAGAGTGGCGCGCTCAGAGGCTCAACACGCTGTGTC 4155
 QY 181 GAGAAAGCCAGAGATCTTCTTATACCCCTGGAGACATGCTCAACCCAAATAAGCC 240
 DB 4156 GAGAAAGCCAGAGATCTTCTTATACCCCTGGAGACATGCTCAACCCAAATAAGCC 4215
 QY 241 TACCAATGCGCTGGGGCTGGCGCTACCTGCAAGCAAGAAATATATCTTCGTACCTG 300
 DB 4216 TACCAATGCGCTGGGGCTGGCGCTACCTGCAAGCAAGAAATATATCTTCGTACCTG 4275
 QY 301 AAGTGGACAAATCTGTGTGTGTGCTTCCCTTGAAGTCAAGAGACATCAATCAAGCTA 360
 DB 4276 AAGTGGACAAATCTGTGTGTGTGCTTCCCTTGAAGTCAAGAGACATCAATCAAGCTA 4335
 QY 361 TCTGACTAGGAGATTTTCAGAGCATATTCATGAGAGGCGCCCTGAGGGTGAAGGCACT 420
 DB 4336 TCTGACTAGGAGATTTTCAGAGCATATTCATGAGAGGCGCCCTGAGGGTGAAGGCACT 4395
 QY 421 CCTGCTACCAAGGCCCGCAGAGATCAGGCGCTGCATGTATATGAGAGAGTAAATATG 480
 DB 4396 CCTGCTACCAAGGCCCGCAGAGATCAGGCGCTGCATGTATATGAGAGAGTAAATATG 4455
 QY 481 TTCCTATGAGATGCTGCTTACAGATTGCTGTAGAGACAGCGCTGCATGAGGCGAC 540
 DB 4456 TTCCTATGAGATGCTGCTTACAGATTGCTGTAGAGACAGCGCTGCATGAGGCGAC 4515
 QY 541 CACCAAGCTCCAGATTGCCAAGAGAGCTGTCCAAAGGCGATCCGCCGCTGTGGGCGAGCCG 600
 DB 4516 CACCAAGCTCCAGATTGCCAAGAGAGCTGTCCAAAGGCGATCCGCCGCTGTGGGCGAGCCG 4575
 QY 601 GAGGAAGTGCATTCGCGGAGCTGCAGGCGCTCATGATGAGAGTGTGGAGACTAAGCA 660
 DB 4576 GAGGAAGTGCATTCGCGGAGCTGCAGGCGCTCATGATGAGAGTGTGGAGACTAAGCA 4635
 QY 661 GAGGAAGTGCATTCGCGGAGCTGCAGGCGCTCATGATGAGAGTGTGGAGACTAAGCA 720
 DB 4636 GAGGAAGTGCATTCGCGGAGCTGCAGGCGCTCATGATGAGAGTGTGGAGACTAAGCA 4695
 QY 721 TTCATGTATGAACTGT 780
 DB 4696 TTCATGTATGAACTGT 4755
 QY 781 TACACCGT 840
 DB 4756 TACACCGT 4815
 QY 841 GAGAAAGGCTCATGAGAGT 900

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPLO
CC at ftp.wiplo.int/pub/published_pcl_sequences.

XX Sequence 2652 BP: 565 A; 824 C; 766 G; 497 T; 0 other;

Query Match 97.6%; Score 2132.4; DB 23; Length 2652;
Best Local Similarity 99.4%; Pred. No. 0;

Matches 2172; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

```
QY 1 ATGCGAGGACCTGGGGCCACCGATGAGAGAACTTCCGAGTCCGGGAGGAG 60
DB 445 ATGCTGAGGACCTGGGGCCACCGATGAGAGAACTTCCGAGTCCGGGAGGAG 504
QY 61 GCCAGATGCTCAGCGCGCTGAGACACCCCTGATCGTGGCGCTATCGGATCAGATC 120
DB 505 GCCAGATGCTCAGCGCGCTGAGACACCCCTGATCGTGGCGCTATCGGATCAGATC 564
QY 121 CACCGGCTGCTGGCCCTGAGAGCTCGCGCTGAGACAGCTTCAACACCGTGTCTCC 180
DB 565 CACCGGCTGCTGGCCCTGAGAGCTCGCGCTGAGACAGCTTCAACACCGTGTCTCC 624
QY 181 GAGAACGCGAGAGATCTCTTTATACCCCTGGAGACATGCTCCCAAAAATAGCC 240
DB 625 GAGAACGCGAGAGATCTCTTTATACCCCTGGAGACATGCTCCCAAAAATAGCC 664
QY 241 TACAGATGCTCGGCGCTGAGCTACGACAAAGAAAATCATCTCTGTGACTG 300
DB 685 TACAGATGCTCGGCGCTGAGCTACGACAAAGAAAATCATCTCTGTGACTG 744
QY 301 AAGTGGAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
DB 745 AAGTGGAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 804
QY 361 TGTGATGAGGAGTTTGAAGGAGTTCATTCATGAGAGGCGCCCTAGGCGTGGAGGACT 420
DB 805 TGTGATGAGGAGTTTGAAGGAGTTCATTCATGAGAGGCGCCCTAGGCGTGGAGGACT 864
QY 421 CCGTGTACAGGCGCCAGAGATCAGGCGCTGCTGATGATGATGATGATGATGATG 480
DB 865 CCGTGTACAGGCGCCAGAGATCAGGCGCTGCTGATGATGATGATGATGATGATG 924
QY 481 TTTCTCTATGAGATGCTGCTCTAGAGTGTCTCTAGAGAGCGCCCTGCTGAGGCGAC 540
DB 925 TTTCTCTATGAGATGCTGCTCTAGAGTGTCTCTAGAGAGCGCCCTGCTGAGGCGAC 984
QY 541 CACCACTCCAGATGTCAGAAAGTGTCCCAAGGGATCCGCGGTTTGGGGAGCGCG 600
DB 985 CACCACTCCAGATGTCAGAAAGTGTCCCAAGGGATCCGCGGTTTGGGGAGCGCG 1044
QY 601 GAGGAAGTCAATCCGCGGAGTCAAGGCGCTGATGATGATGATGATGATGATGATG 660
DB 1045 GAGGAAGTCAATCCGCGGAGTCAAGGCGCTGATGATGATGATGATGATGATGATG 1104
QY 661 GAGGAAGTCAATCCGCGGAGTCAAGGCGCTGATGATGATGATGATGATGATGATG 720
DB 1105 GAGGAAGTCAATCCGCGGAGTCAAGGCGCTGATGATGATGATGATGATGATGATG 1164
QY 721 TTTCTCTATGAGATGCTGCTCTAGAGTGTCTCTAGAGAGCGCCCTGCTGAGGCGAC 780
DB 1165 TTTCTCTATGAGATGCTGCTCTAGAGTGTCTCTAGAGAGCGCCCTGCTGAGGCGAC 1224
QY 781 TACACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 1225 TACACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1284
QY 841 GAGAGGCGCTCATGAGAGTGCAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 900
```

```
DB 1285 GAGAAAGGCGCTCATGAGAGTGCAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 1344
QY 901 CTCAGATCCAGAGATCCCTGTGTGACAGACCGAGAGACAGAAAATCTATCTACAC 960
DB 1345 CTCAGATCCAGAGATCCCTGTGTGACAGACCGAGAGACAGAAAATCTATCTACAC 1404
QY 961 CTCAGAGGATGTGCGCCCTTAAACACACCCCAAGGCGTGTGTGTGTGTGTGTGTGT 1020
DB 1405 CTCAGAGGATGTGCGCCCTTAAACACACCCCAAGGCGTGTGTGTGTGTGTGTGTGT 1464
QY 1021 ACCCTGCTTGTGCGGCTGCTGCTTAAAGAAATCTTACTGCTGTGCTGCTGCTGCTG 1080
DB 1465 ACCCTGCTTGTGCGGCTGCTGCTTAAAGAAATCTTACTGCTGTGCTGCTGCTGCTG 1524
QY 1081 GCCGATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
DB 1525 GCCGATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1584
QY 1141 TACCTGTGCTCAGACACACAGCAAGAGTCCAGATTGATGATGATGATGATGATGATG 1200
DB 1585 TACCTGTGCTCAGACACACAGCAAGAGTCCAGATTGATGATGATGATGATGATGATG 1644
QY 1201 CAGAACCCCTACCCAGTGAAGGAGTGAAGTGTGTCAGAGGCGCTGAGGCTGTGATC 1260
DB 1645 CAGAACCCCTACCCAGTGAAGGAGTGAAGTGTGTCAGAGGCGCTGAGGCTGTGATC 1704
QY 1261 AGCAATGAGGCGCGGCTCTCTTGTATGATGATGATGATGATGATGATGATGATGATG 1320
DB 1705 AGCAATGAGGCGCGGCTCTCTTGTATGATGATGATGATGATGATGATGATGATGATG 1764
QY 1321 GAGCCTTACATGAGGCGGCTCTCTTGTATGATGATGATGATGATGATGATGATGATG 1380
DB 1765 GAGCCTTACATGAGGCGGCTCTCTTGTATGATGATGATGATGATGATGATGATGATG 1824
QY 1381 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
DB 1825 GAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1884
QY 1441 ACCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
DB 1885 ACCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1944
QY 1501 CCGGTGCGGCGCTTGTGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1945 CCGGTGCGGCGCTTGTGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2004
QY 1561 CCGTGAAGGAGGAGTCCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
DB 2005 CCGTGAAGGAGGAGTCCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2064
QY 1621 ATCTGATCCAGAGAGATCACTGACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 2065 ATCTGATCCAGAGAGATCACTGACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2124
QY 1681 CCACCCCGGAGGCTGCGAGAGTCCCTCAAGCTTCCCAAGCTTCCCAAGCACTTTC 1740
DB 2125 CCACCCCGGAGGCTGCGAGAGTCCCTCAAGCTTCCCAAGCTTCCCAAGCACTTTC 2184
QY 1741 AGTGTGCTTGTTCACACGAGTCCGAGAGTCCAGAGATGATGATGATGATGATGATG 1800
DB 2185 AGTGTGCTTGTTCACACGAGTCCGAGAGTCCAGAGATGATGATGATGATGATGATG 2244
QY 1801 TCCGACAGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1859
DB 2245 TCCGACAGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2304
QY 1860 GAGGCGCGTGAAG-ATCTTGTGCGGCTGAGAGAGCTCATTTGGTCCCAAGGCGGCTG 1918
DB 2305 GAGGCGCGGGAAGAAATCTGCGGCTGAGAGAGCTCATTTGGTCCCAAGGCGGCTG 2364
QY 1919 ATGTTATGCTCATTTGCGGCTGAGAGAGTCTCAAGGCGGAGGCGGAGGCTGATG 1978
DB 2365 ATGTTATGCTCATTTGCGGCTGAGAGAGTCTCAAGGCGGAGGCGGAGGCTGATG 2424
```



```

Db 854 TACACCTGCTGTTTGGAGATGAAAAAGAGATCCAGAACTACAGGTGTGTAACA 913
QY 841 GAGAAAGGCTCATGAGATGAGATGATGCTCCCTGGATGAAGTGAAGTCCAG 900
Db 914 GAGAAAGGCTCATGAGATGAGATGATGCTCCCTGGATGAAGTGAAGTCCAG 973
QY 901 CTCGAGGTCAGAGATGCTGTGAGACGACCGAGAGACAGAAAAATCTACATCTACAC 960
Db 974 CTCGAGGTCAGAGATGCTGTGAGACGACCGAGAGACAGAAAAATCTACATCTACAC 1033
QY 961 CTCGAGGTCAGAGATGCTGTGAGACGACCGAGAGATGCTGTGAGATGCTGTGAG 1020
Db 1034 CTCGAGGTCAGAGATGCTGTGAGACGACCGAGAGATGCTGTGAGATGCTGTGAG 1093
QY 1021 ACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1094 ACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1153
QY 1081 GCCGATGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1140
Db 1154 GCCGATGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1213
QY 1141 TACCTGTGCTACACACACACACACACACACACACACACACACACACACACACAC 1200
Db 1214 TACCTGTGCTACACACACACACACACACACACACACACACACACACACACACAC 1273
QY 1201 CAGAACCCCTACACAGTGAAGGTCATGAGAGGTCATGAGAGGTCATGAGAGG 1260
Db 1274 CAGAACCCCTACACAGTGAAGGTCATGAGAGGTCATGAGAGGTCATGAGAGG 1333
QY 1261 AGCAATGGGCGCGGCGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1320
Db 1334 AGCAATGGGCGCGGCGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1393
QY 1321 GAGCCCTACATGAGGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1380
Db 1394 GAGCCCTACATGAGGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1453
QY 1381 GAGGAGGTCGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1440
Db 1454 GAGGAGGTCGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1513
QY 1441 ACCTACACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1500
Db 1514 ACCTACACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1573
QY 1501 CCCGTCGCGCGCTGTGAGACAGGACCGGACCGACCGACCGACCGACCGACCGAC 1560
Db 1574 CCCGTCGCGCGCTGTGAGACAGGACCGGACCGGACCGACCGACCGACCGACCG 1633
QY 1561 CCTGAGGAGGACATGATGCGGAGAGGATGATGATGATGATGATGATGATGATGATG 1620
Db 1634 CCTGAGGAGGACATGATGCGGAGAGGATGATGATGATGATGATGATGATGATGATG 1693
QY 1621 ATCTGATGTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1694 ATCTGATGTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1753
QY 1681 CCACCCCGGACGCTGCGAGGTCGCCCTGAGCCTGCCAGCTGCCAGAGCTGCTTC 1740
Db 1754 CCACCCCGGACGCTGCGAGGTCGCCCTGAGCCTGCCAGCTGCCAGAGCTGCTTC 1813
QY 1741 ACTGAGCTTTTCTCAGCAGCTGCGAGGATGATGATGATGATGATGATGATGATGAT 1800
Db 1814 ACTGAGCTTTTCTCAGCAGCTGCGAGGATGATGATGATGATGATGATGATGATGAT 1873
QY 1801 TCCGAGAGGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1874 TCCGAGAGGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1933
QY 1861 CAGGCGCTGTAAGATGCTGCGCTGCAAGACCTCATTTGGTCCCGACGCGCTGAGAT 1933

```

```

QY 1921 GTATGCTATTTGGCTGAGAGATTTCTGAGCCAGCGGCGCGAGTCAATGCCGTC 1980
Db 1994 GTATGCTATTTGGCTGAGAGATTTCTGAGCCAGCGGCGCGAGTCAATGCCGTC 2053
QY 1981 TTTAAAGCCCGAGAGCTACTCCGATGGGTGCTGTGATGCTGCGGTGGCAAG 2040
Db 2054 TTTAAAGCCCGAGAGCTACTCCGATGGGTGCTGTGATGCTGCGGTGGCAAG 2113
QY 2041 GACACTGTTGTGACCTTTGAAAAATGAAACACAGAGTGGTGGCTGGAGG 2100
Db 2114 GACACTGTTGTGACCTTTGAAAAATGAAACACAGAGTGGTGGCTGGAGG 2173
QY 2101 GCGTGGGCGCGAGGAGTGTGACATTTTCTACAGTGTGAGAGAGTGGCGGCTG 2160
Db 2174 GCGTGGGCGCGAGGAGTGTGACATTTTCTACAGTGTGAGAGAGTGGCGGCTG 2233
QY 2161 GAGGCTGCACTCGCAAGAGAGG 2184
Db 2234 GAGGCTGCACTCGCAAGAGAGG 2257

RESULT 4
AA572238
ID AA572238 standard; cDNA; 2652 BP.
XX
XX AA572238;
AC
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #8042.
DE
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PE
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI: 2001-639362/73.
DR P-FSDB; ABG08051.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID No 8042; 103bp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromo-
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in

```

CC	AD33695		
ID	AD33695	standard; cDNA; 3564 BP.	
XX			
AC			
XX	AD33695;		
DT			
XX			
DT	01-JUL-2002	(first entry)	
XX			
XX	Human secreted protein-encoding gene 4 cDNA clone HDDM90. SEQ ID NO:14.		
DE			
KW	Human; secreted protein; immune disorder; antiallergic; anti-rheumatic;		
KW	rheumatoid arthritis; breast neoplasia; breast cancer; antiarthritic;		
KW	neurological disease; Alzheimer's disease; Parkinson's disease; trauma;		
KW	Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania		
KW	antihistaminatory; ophthalmological; dermatological; immunostimulatory;		
KW	immunomodulatory; immunosuppressive; antibacterial; antiparasitic;		
KW	gene therapy; autoimmune disease; Huntington's disease; meningitis;		
KW	demyelinating disease; peripheral neuropathy; congenital malformation;		
KW	spinal cord injury; peripheral neuropathy; ischemia; perception;		
KW	multiple sclerosis; infarction; haemorrhage; schizophrenia; dementia;		
KW	depression; panic disorder; learning disability; AIDS; feeding disorder;		
KW	hyperproliferative disorder; sleep pattern; cardiovascular disorder;		
KW	reproductive disorder; digestive system disorder; behavioural disorder;		
KW	gene; ss.		
OS			
XX	Homo sapiens.		
FT			
FT	Key	Location/Qualifiers	
FT	CDS	140..2260	
FT		/+tag- a	
FT		/product- "Human secreted protein"	
FT	sig_peptide	140..211	
FT		/+tag- b	
FT	mat_peptide	212..2257	
FT		/+tag- c	
FT		/product- "Human mature secreted protein"	
PN			
XX	MO200216390-A1.		
PD			
XX	28-FEB-2002.		
XX			
PF	17-JAN-2001; 2001WO-US01435.		
XX			
PR	18-AUG-2000; 2000US-226282P.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
P1	Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;		
P1	Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M,		
P1	Ni J;		
XX			
DR	WP1: 2002-304113/34.		
DR	P-PSDB; AAE21194.		
XX			
XX	An isolated nucleic acid molecule (I) comprising a polynucleotide which		
PT	encodes a polypeptide useful in the diagnosis and treatment of		
PT	disorders e.g. immune disorders -		
XX			
PS	Claim 1; Page 442-443; 534pp; English.		
XX			
CC	AD33692-AD33736 represent cDNAs corresponding to 21 human secreted		
CC	protein genes, and AAE21191-AAE21235 represent the proteins they encode.		
CC	AAE21236-AAE21280 represent human secreted protein fragments. The genes		
CC	and their corresponding secreted proteins are useful for preventing,		
CC	treating or ameliorating medical conditions, e.g., by protein or gene		
CC	therapy. Pathological conditions can be diagnosed by determining the		
CC	amount of the new protein in a sample or by determining the presence of		
CC	mutations in the new genes. Specific uses are described for each of the		
CC	21 genes, based on the tissues in which they are most highly expressed,		
CC	and include developing products for the diagnosis or treatment of		
CC	immune or autoimmune diseases e.g. AIDS (acquired immune deficiency		
CC	syndrome), asthma, anemia and rheumatoid arthritis, breast neoplasia		

Query Match	99.9%	Score 2180.8	DB 24	Length 3564
Best Local Similarity	99.9%	Pred. No. 0		
Matches 2182	Conservative	0	Mismatches	2
			Indels	0
			Gaps	
CC and breast cancer, neurological diseases e.g. Alzheimer's disease,				
CC Parkinson's disease, Huntington's disease, Tourette syndrome,				
CC meningitis, demyelinating disease, peripheral neuropathies, neoplasia				
CC trismus, congenital malformations, spinal cord injuries, toxic				
CC neuropathies induced by neurotoxins, peripheral neuropathies, multipl				
CC sclerosis, ischemia and infection, haemorrhages, schizophrenia, man				
CC dementia, depression, panic disorder, learning disabilities, ALS,				
CC altered behaviours e.g. disorders in feeding, sleep patterns, balance				
CC and perception, encephalitis, disorders in cardiovascular, neural/				
CC sensory, reproductive and digestive systems, behavioural disorders and				
CC hyperproliferative disorder. The present sequence represents a human				
CC secreted protein-encoding cDNA of the invention.				
XX				
Sequence 3564 BP; 844 A; 1014 C; 958 G; 748 T; 0 other:				
Query Match	99.9%	Score 2180.8	DB 24	Length 3564
Best Local Similarity	99.9%	Pred. No. 0		
Matches 2182	Conservative	0	Mismatches	2
			Indels	0
			Gaps	
QY 1 ATGCTGAGGACACCTGGCGGGCCACCGATGACATGAAGACTTCTCCGAGTTCCGGCAGAG 60				
DB 74 ATGTGTAGGACACCTGGCGGGCCACCGATGACATGAAGACTTCTCCGAGTTCCGGCAGAG 133				
QY 61 GCCAGCATGCTGACACGGCGCTGACACACCCCTGCATGCTGGCGCTCATTCGSCATCAGCATC 120				
DB 134 GCCAGCATGCTGACACGGCGCTGACACACCCCTGCATGCTGGCGCTCATTCGSCATCAGCATC 193				
QY 121 CACCCGCTCTGCTTGGCCCTGAGACTGCGCGCTCAGCAGCCTCAACACCGTGTCTCC 180				
DB 194 CACCCGCTCTGCTTGGCCCTGAGACTGCGCGCTCAGCAGCCTCAACACCGTGTCTCC 253				
QY 181 GAGAACGCCAGAGATTCTCTCTTTATACCCCTGGGACATGCTCACCACCAAAATAATAGCC 240				
DB 254 GAGAACGCCAGAGATTCTCTCTTTATACCCCTGGGACATGCTCACCACCAAAATAATAGCC 313				
QY 241 TACCAAGATCGCCTGGGGCGTGGCCCTACTCTGACAAAGAAAACATCATCTCTGTGACCTG 300				
DB 314 TACCAAGATCGCCTGGGGCGTGGCCCTACTCTGACAAAGAAAACATCATCTCTGTGACCTG 373				
QY 301 AAGTGGACAACATTCTGTGTGTGTGCTCCCTTGACCTGACGAGAGGACATCAACATCAAGCTA 360				
DB 374 AAGTGGACAACATTCTGTGTGTGTGCTCCCTTGACCTGACGAGAGGACATCAACATCAAGCTA 433				
QY 361 TCTGACTACGGGATTTGGAGGCACTATTCCATGAGGGCGCCTTAGCGGTGAGGGCACT 420				
DB 434 TCTGACTACGGGATTTGGAGGCACTATTCCATGAGGGCGCCTTAGCGGTGAGGGCACT 493				
QY 421 CCTGCTTCCAGGGCCCAAGAGATAGGCGCTGCAATGTATATATAGAGAGATAGATAG 480				
DB 494 CCTGCTTCCAGGGCCCAAGAGATAGGCGCTGCAATGTATATATAGAGAGATAGATAG 553				
QY 481 TTCTCCTATGGAATGTGTGCTTACGAAATTCCTGTGACAGAGAGGCCCTGACATGGGCCAC 540				
DB 554 TTCTCCTATGGAATGTGTGCTTACGAAATTCCTGTGACAGAGAGGCCCTGACATGGGCCAC 613				
QY 541 CACCACTCCAGATTGCCAAGAGCTGTCCAAAGGGCATCCGCCGGTTCTGGGGGACGGC 600				
DB 614 CACCACTCCAGATTGCCAAGAGCTGTCCAAAGGGCATCCGCCGGTTCTGGGGGACGGC 673				
QY 601 GAGGAAGTGCAGTTTCCGGGCACCTCAGAGCGTCAATGATGAGAGTGTGGACACTAAGCCA 660				
DB 674 GAGGAAGTGCAGTTTCCGGGCACCTCAGAGCGTCAATGATGAGAGTGTGGACACTAAGCCA 733				
QY 661 GAGGAAGTGCAGTTTCCGGGCACCTCAGAGCGTCAATGATGAGAGTGTGGACACTAAGCCA 720				
DB 734 GAGGAAGTGCAGTTTCCGGGCACCTCAGAGCGTCAATGATGAGAGTGTGGACACTAAGCCA 793				
QY 721 TTCATGTATGAATGTGTGTGTGGAGACAGACACTTCTTCTCATCCAGGGCCAGAG 780				
DB 794 TTCATGTATGAATGTGTGTGTGGAGACAGACACTTCTTCTCATCCAGGGCCAGAG 853				
QY 781 TACACCGTGTGTGTGTGGAGTGAAGAAAGAGAGTCCAGAACTACACGGTGTGAACACA 840				

Db 68 ATGCTAGAGGACACCTGCGGCGCACCCGATGAGAAATCTCCGAGTTCGCGAGGAG 127
Oy 61 GCCAGATGCTGACAGCGGTGAGAGACCCCTGATGCGGCTCATGGGATCAGCATC 120
Db 128 GCCAGATGCTGACAGCGGTGAGAGACCCCTGATGCGGCTCATGGGATCAGCATC 187
Oy 121 CACCGCTCTGCTGCGGCTGAGAGTCCGCGCTCAGAGGCTCAACACCGCTGCTGC 180
Db 188 CACCGCTCTGCTGCGGCTGAGAGTCCGCGCTCAGAGGCTCAACACCGCTGCTGC 247
Oy 181 GAGAACGCCAGAGATCTTCTCTTATACCCCTGAGACATCTCAACCAAAATAGCC 240
Db 248 GAGAACGCCAGAGATCTTCTCTTATACCCCTGAGACATCTCAACCAAAATAGCC 307
Oy 241 TACAGATGCTGCGGCGCTGAGCTACCTGACAGAAAGAAATCATCTCTGAGCTG 300
Db 308 TACAGATGCTGCGGCGCTGAGCTACCTGACAGAAAGAAATCATCTCTGAGCTG 367
Oy 301 AAGTCGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 368 AAGTCGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
Oy 361 TCTGATCGGAGATTTGAGGAGATCATTCATGAGGAGGCGGCTGAGGAGGACT 420
Db 428 TCTGATCGGAGATTTGAGGAGATCATTCATGAGGAGGCGGCTGAGGAGGACT 487
Oy 421 CCGTGCATCAGAGGCGGCTGAGAGATCAGGCTGCTGATGATGAGAGGATAGT 480
Db 488 CCGTGCATCAGAGGCGGCTGAGAGATCAGGCTGCTGATGATGAGAGGATAGT 547
Oy 481 TCTCTCTATGAGAGTGTGCTTACAGATTTGCTGACAGAGACGCGCTGAGCTG 540
Db 548 TCTCTCTATGAGAGTGTGCTTACAGATTTGCTGACAGAGACGCGCTGAGCTG 607
Oy 541 CACCGATCCAGATTTGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 608 CACCGATCCAGATTTGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 667
Oy 601 GAGAGATGCAATTTCCGCGGCTGAGAGGCTCATGATGATGATGATGATGATG 660
Db 668 GAGAGATGCAATTTCCGCGGCTGAGAGGCTCATGATGATGATGATGATGATG 727
Oy 661 GAGAGATGCAATTTCCGCGGCTGAGAGGCTCATGATGATGATGATGATGATG 720
Db 728 GAGAGATGCAATTTCCGCGGCTGAGAGGCTCATGATGATGATGATGATGATG 787
Oy 721 TTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 788 TTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 847
Oy 781 TACACCGTGTGTTTGGAGTGAAGAAAGAGAGTCCAGAGATCAGAGTGTGAACA 840
Db 848 TACACCGTGTGTTTGGAGTGAAGAAAGAGAGTCCAGAGATCAGAGTGTGAACA 907
Oy 841 GAGAAAGGCTCATGAGAGTGAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 908 GAGAAAGGCTCATGAGAGTGAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 967
Oy 901 CTTCCAGGTCAGAGATCCTGCTGAGACAGCAGAGAGACCAAAATCTACATCTAC 960
Db 968 CTTCCAGGTCAGAGATCCTGCTGAGACAGCAGAGAGACCAAAATCTACATCTAC 1027
Oy 961 CTTCCAGGTCAGAGATCCTGCTGAGACAGCAGAGAGACCAAAATCTACATCTAC 1020
Db 1028 CTTCCAGGTCAGAGATCCTGCTGAGACAGCAGAGAGACCAAAATCTACATCTAC 1087
Oy 1021 ACCTGCTTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1088 ACCTGCTTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1147
Oy 1081 GCGGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1148 GCGGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207

Oy 1141 TACCTGTGCTACACAGACCCAGATGCTCAAGTTCACATGCGGATGAAGACGACG 1200
Db 1208 TACCTGTGCTACACAGACCCAGATGCTCAAGTTCACATGCGGATGAAGACGACG 1267
Oy 1201 CAGAACCCCTACAGATGAGAGGCTGATGATGATGATGATGATGATGATGATG 1260
Db 1268 CAGAACCCCTACAGATGAGAGGCTGATGATGATGATGATGATGATGATGATG 1327
Oy 1261 AGCATGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1328 AGCATGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1387
Oy 1321 GAGCCCTACATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Db 1388 GAGCCCTACATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1447
Oy 1381 GAGAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
Db 1448 GAGAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1507
Oy 1441 ACCTACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Db 1508 ACCTACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1567
Oy 1501 CCGTGCAGGCTTGTGAGACAGGACCCCGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Db 1568 CCGTGCAGGCTTGTGAGACAGGACCCCGGAGGAGGAGGAGGAGGAGGAGGAG 1627
Oy 1561 CCGTGCAGGCTTGTGAGACAGGAGCTGAGACATGATGATGATGATGATGATG 1620
Db 1628 CCGTGCAGGCTTGTGAGACAGGAGCTGAGACATGATGATGATGATGATGATG 1687
Oy 1621 ATCTGATCCACAGATGATGATGATGATGATGATGATGATGATGATGATG 1680
Db 1688 ATCTGATCCACAGATGATGATGATGATGATGATGATGATGATGATGATG 1747
Oy 1681 CCACCCGCGAGAGTGTGAGAGTCCCTCAAGCTCCAGCTCCCAAGAGTCTTC 1740
Db 1748 CCACCCGCGAGAGTGTGAGAGTCCCTCAAGCTCCAGCTCCCAAGAGTCTTC 1807
Oy 1741 AGTGTGCTTCTCCACAGAGTGTGAGAGTCCCTCAAGCTCCCAAGAGTCTTC 1800
Db 1808 AGTGTGCTTCTCCACAGAGTGTGAGAGTCCCTCAAGCTCCCAAGAGTCTTC 1867
Oy 1801 TCCGAGGTCGAGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 1868 TCCGAGGTCGAGATGATGATGATGATGATGATGATGATGATGATGATG 1927
Oy 1861 CAGGCGGTGAGATCCTGCGCTCAGAGACTCATTTGGTCCCAAGGCGGCTGAGAT 1920
Db 1928 CAGGCGGTGAGATCCTGCGCTCAGAGACTCATTTGGTCCCAAGGCGGCTGAGAT 1987
Oy 1921 GTTATGCTATTTGCGCTGAGAGATTTGATGATGATGATGATGATGATGATG 1980
Db 1988 GTTATGCTATTTGCGCTGAGAGATTTGATGATGATGATGATGATGATGATG 2047
Oy 1981 TTAAGAGCCGAGAGTCACTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 2040
Db 2048 TTAAGAGCCGAGAGTCACTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2107
Oy 2041 GACACTGTGCTGAGACTTGTGAAATGAAACACAGAGTGTGCTGCTGCTGAGAG 2100
Db 2108 GACACTGTGCTGAGACTTGTGAAATGAAACACAGAGTGTGCTGCTGCTGAGAG 2167
Oy 2101 GCGTGGGCGGAGGAGTGTGACATTTTCTACAGTCTTACAGAGAGTGGCGGCTG 2160
Db 2168 GCGTGGGCGGAGGAGTGTGACATTTTCTACAGTCTTACAGAGAGTGGCGGCTG 2227
Oy 2161 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2184
Db 2228 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2251

QY 1441 ACCTACAGCTGTGTGCTCCGCTACTTCTGCGGGGTCCCGAGCCCTCAGGACATGTTT 1500
 |||||||
 DB 1462 ACCTACAGCTGTGTGCTCCGCTACTTCTGCGGGGTCCCGAGCCCTCAGGACATGTTT 1521
 |||||||
 QY 1501 CCCGTGGCCCTTGGACAGGAAACCCCGGAGCCAGCCACAGGCCAACCCAAAGGTG 1560
 |||||||
 DB 1522 CCCGTGGCCCTTGGACAGGAAACCCCGGAGCCAGCCACAGGCCAACCCAAAGGTG 1581
 |||||||
 QY 1561 CCTGAGGGGGAGTCCATGCGGAGAGTGCATGATGATGAGAGAGAGTGGGACAGAG 1620
 |||||||
 DB 1582 CCTGAGGGGGAGTCCATGCGGAGAGTGCATGATGATGAGAGAGAGTGGGACAGAG 1641
 |||||||
 QY 1621 ATCCGATCCACAGCAAGATCACTACTACTCTCTCATCTCTCTCTCTCTCTCTCTCT 1680
 |||||||
 DB 1642 ATCCGATCCACAGCAAGATCACTACTACTCTCTCATCTCTCTCTCTCTCTCTCTCT 1701
 |||||||
 QY 1681 CCACCCCGCCAGGCTGCGAGGTCCCTCTCAAGCTCCAGCTCCCGAGCAAGTCTCTCC 1740
 |||||||
 DB 1702 CCACCCCGCCAGGCTGCGAGGTCCCTCTCAAGCTCCAGCTCCCGAGCAAGTCTCTCC 1761
 |||||||
 QY 1741 AGTGTGCTTTCTCCACGCTGCGAGAGTCTGACATGCTATGCTATGCTGCTGCTGCTG 1800
 |||||||
 DB 1762 AGTGTGCTTTCTCCACGCTGCGAGAGTCTGACATGCTATGCTATGCTGCTGCTGCTG 1821
 |||||||
 QY 1801 TCCGACAGGCTGAGCATGACCTGACCCCATGAGCGGGGAGAGCTTGACGACGACCTG 1860
 |||||||
 DB 1822 TCCGACAGGCTGAGCATGACCTGACCCCATGAGCGGGGAGAGCTTGACGACGACCTG 1881
 |||||||
 QY 1861 CAGGCGGTGAAGATCTCTGCGCTGAGAGACTCTATTTGGGTCCCGAGCGCGGTGAGAT 1920
 |||||||
 DB 1882 CAGGCGGTGAAGATCTCTGCGCTGAGAGACTCTATTTGGGTCCCGAGCGCGGTGAGAT 1941
 |||||||
 QY 1921 GTTATCGCTATGCTGCTGAGAGAGATCTGTAAGCCCGAGCGGGGCTGATCTGCTGCTG 1980
 |||||||
 DB 1942 GTTATCGCTATGCTGCTGAGAGAGATCTGTAAGCCCGAGCGGGGCTGATCTGCTGCTG 2001
 |||||||
 QY 1981 TTAAGAAGCCGAGAGCTGCTCGCATGGGGTGTGTGTGATGCTGCTGCTGCTGCTGCTG 2040
 |||||||
 DB 2002 TTAAGAAGCCGAGAGCTGCTCGCATGGGGTGTGTGTGATGCTGCTGCTGCTGCTGCTG 2061
 |||||||
 QY 2041 GACACTGTTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGCTGCTGCTGCTGCTG 2100
 |||||||
 DB 2062 GACACTGTTGTGTGACCTTTGAAATGAAACACAGAGTGTGCTGCTGCTGCTGCTGCTG 2121
 |||||||
 QY 2101 GGTCTGGGGGCGGAGGAGTGTGACATTTCTACAGTCTCTAGAGAGTGTGGGCGGCTG 2160
 |||||||
 DB 2122 GGTCTGGGGGCGGAGGAGTGTGACATTTCTACAGTCTCTAGAGAGTGTGGGCGGCTG 2181
 |||||||
 QY 2161 GAGGCTTGACCTGCGACAGAGAGG 2184
 |||||||
 DB 2182 GAGGCTTGACCTGCGACAGAGAGG 2205
 |||||||

RESULT 2

AAD33718
 ID AAD33718 standard; cDNA; 3579 BP.

AC AAD33718;

DT 01-JUL-2002 (first entry)

DE Human secreted protein-encoding gene 4 cDNA clone HDDMW90, SEQ ID NO:37.

XX Human, secreted protein; immune disorder; antiallergic; antihemetic;
 XX Rheumatoid arthritis; breast neoplasia; breast cancer; antiallergic;
 XX neurologic disease; Alzheimer's disease; Parkinson's disease; trauma;
 XX Tourette syndrome; encephalitis; cytostatic; haemostatic; anaemia; mania;
 XX antiinflammatory; ophthalmological; dermatological; immunostimulatory;
 XX immunomodulatory; immunosuppressive; antibacterial; antiparasitic;
 XX gene therapy; autoimmune disease; Huntington's disease; meningitis;
 XX demyelinating disease; peripheral neuropathy; congenital malformation;
 XX spinal cord injury; peripheral neuropathy; ischaemia; perception;
 XX multiple sclerosis; infarction; haemorrhage; schizophrenia; dementia;

KW depression; panic disorder; learning disability; ALS; feeding disorder;
 KW hyperproliferative disorder; sleep pattern; cardiovascular disorder;
 KW reproductive disorder; digestive system disorder; behavioural disorder;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 134..2254
 FT /*tag= a
 FT /product= "Human secreted protein"
 FT s1g_peptide 134..205
 FT /*tag= b
 FT mat_peptide 206..2251
 FT /*tag= c
 FT /product= "Mature secreted protein"
 FT
 PN WO200216390-A1.
 XX
 PD 28-FEB-2002.
 XX
 PF 17-JAN-2001; 2001WO-US01435.
 XX
 PR 18-AUG-2000; 2000US-226282P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Komatsu GA, Baker KP, Bire CE, Soppet DR, Olsen HS;
 PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fischella M;
 PI Ni J;
 XX
 DR WPI: 2002-304113/34.
 DR P-PSDB: AAE21217.
 XX
 PT An isolated nucleic acid molecule (1) comprising a polynucleotide which
 PT encodes a polypeptide useful in the diagnosis and treatment of
 PT disorders e.g. immune disorders
 XX
 PS Claim 1: Page 457-458; 534pp; English.
 XX
 CC AAD33692-AAD33736 represent cDNAs corresponding to 21 human secreted
 CC protein genes, and AAE21191-AAE21235 represent the proteins they encode.
 CC AAE21236-AAE21280 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 21 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC immune or autoimmune diseases e.g. AIDS (acquired immune deficiency
 CC syndrome), asthma, anaemia and rheumatoid arthritis, breast neoplasia
 CC and breast cancer, neurological diseases e.g. Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, Tourette syndrome,
 CC meningitis, demyelinating disease, peripheral neuropathies, neoplasia,
 CC trauma, congenital malformations, spinal cord injuries, toxic
 CC neuropathies induced by neurotoxins, peripheral neuropathies, multiple
 CC sclerosis, ischaemia and infarction, haemorrhages, schizophrenia, mania,
 CC dementia, depression, panic disorder, learning disabilities, ALS,
 CC altered behaviours e.g. disorders in feeding, sleep patterns, balance
 CC and perception, encephalitis, disorders in cardiovascular, neural/
 CC sensory, reproductive and digestive systems, behavioural disorders and
 CC hyperproliferative disorder. The present sequence represents a human
 CC secreted protein-encoding cDNA of the invention.
 XX
 SQ Sequence 3579 BP; 834 A; 1021 C; 970 G; 753 T; 1 other:
 Query Match 100.0%; Score 2184; DB 24; Length 3579;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGCTGAGCAGCTGCGGCGCAGATGATGAGAACTTCTCGAGTTCGGGAGAGG*60
 |||||||

XX 12-OCT-2000; 2000OWO-US28066.
PR 15-OCT-1999; 99US-0159542.
PR 17-NOV-1999; 99US-0159514.
PR 14-MAR-2000; 2000OUS-01589027.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Ruben SM, Shi Y, Young PE, Ni J;
XX WPI; 2001-290976/30.
DR P-PSDB; AAE00669.
XX New isolated nucleic acid molecule for producing human protein tyrosine kinase receptor
PT
FT
PS
PS
XX
XX
XX
The present cDNA sequence encodes human protein tyrosine kinase receptor (PTK) from clone HDPSB88.
CC The invention relates to human protein tyrosine kinase (PTK) receptors and their corresponding cDNA molecules. PTK receptors are useful in providing immunological probes for differential identification of the tissues or cell types present in a biological sample. PTK is used in methods for the diagnosis, prevention and treatment of various disorders related to PTK such as immune system disorders (severe combined immunodeficiency (SCID), inflammation); hyperproliferative disorders (neoplasia, sarcoidosis); cardiovascular disorders (arrhythmia, atherosclerosis); central nervous system disorders (multiple sclerosis); neurodegenerative disorders (Alzheimer's disease, Parkinson's disease); blood coagulation disorders (thrombocytopenia); autoimmune disorders (biliary cirrhosis, Crohn's disease); respiratory disorders (asthma, allergy); gastrointestinal disorders (inflammatory bowel disease); cerebrovascular disorders (thrombosis, vascular dementia); brain disorders (phenylketonuria); cancers (such as ovarian, lung, bladder, liver, breast and lymphomas); behavioural disorders (Tourette's syndrome); musculo-skeletal disorders (arthritis, trauma, tendonitis); renal disorders (nephrotic syndrome, glomerulonephritis); metastases of malignancies and related disorders (leukaemia, multiple myeloma); and infections caused by bacteria, viruses, fungi and parasites. PTK is also useful for screening therapeutic compounds. PTK is used as an antigen in a vaccine to raise an immune response against infectious disease. PTK nucleic acids are useful in gene therapy.

Sequence 3496 BP; 812 A; 998 C; 949 G; 737 T; 0 other;

Query Match 100.0%; Score 2184; DB 22; Length 3496;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2184; conservative 0; mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGAGGACACTGTGGGGCACCACCATGCATGAAGAATTTCGAGTTCCGGAGAGG 60
Db 22 ATGCGAGGACACTGTGGGGCACCACCATGCATGAAGAATTTCGAGTTCCGGAGAGG 81

QY 61 GCCAGCATGCTGCACAGCGCTGCAGACACCCCTGCATCGTGCGCTCATGGCATCAGCATC 120
Db 82 GCCAGCATGCTGCACAGCGCTGCAGACACCCCTGCATCGTGCGCTCATGGCATCAGCATC 141

QY 121 CACCGGCTGTGTTGGCCCTGAGAGTCGCCGCGCTGACAGAGGCTCAACAACCGTGTGCC 180
Db 142 CACCGGCTGTGTTGGCCCTGAGAGTCGCCGCGCTGACAGAGGCTCAACAACCGTGTGCC 201

QY 181 GAGAACGCCAGAGATTCTTCCTTTACCCCTGGGACACATGCTCACCCAAAAAATAGCC 240
Db 202 GAGAACGCCAGAGATTCTTCCTTTACCCCTGGGACACATGCTCACCCAAAAAATAGCC 261

QY 241 TAACGATGCGCTTGGGCTTGGCTTAACCTGCACAAGAAAACATCATCTTCTGTGACCTG 300
Db 262 TAACGATGCGCTTGGGCTTGGCTTAACCTGCACAAGAAAACATCATCTTCTGTGACCTG 321

QY 361 AAGTGGACAAACATTCGTGTGTGTGTCCTTGAGCTCAAGAGCACAATCAATCAGTCAAG 360
Db 381 AAGTGGACAAACATTCGTGTGTGTGTCCTTGAGCTCAAGAGCACAATCAATCAATCAAGTCA 381

OY	361	TCCTACATACGGGATTTTGAGACAGTCATCTTCATATGAGGGCCCTTAGCGTGGAGGGCACT	420
Db	382	TCCTACTACGGGATTTTGAGGCACTCATCTTCATATGAGGGCCCTTAGCGTGGAGGGCACT	441
OY	421	CCCTGGCTACACAGGCCCCAGAGATCCAGGCGCTCGCATTTGTATATGATGAGAAGGTAGATATG	480
Db	442	CCCTGGCTACACAGGCCCCAGAGATCCAGGCGCTCGCATTTGTATATGATGAGAAGGTAGATATG	501
OY	481	TTCTCCATATGGAATGGTGGCTCTACGAGTCTCTCAAGACAGGCCCTGCACTGGGCCAC	540
Db	502	TTCTCCATATGGAAGGTGGCTCTACGAGTCTCTCAAGACAGGCCCTGCACTGGGCCAC	561
OY	541	CACCAAGCTCCAGATTTGGCCAAGAAAGCTCTCCAAAGGCTCCCGGCCGGCTGTCGGGGGACGGC	600
Db	562	CACCAAGCTCCAGATTTGGCCAAGAAAGCTCTCCAAAGGCTCTCCCGGCCGGCTGTCGGGGGACGGC	621
OY	601	GAGGAAGTGCAGTTCGCGGCACTGCAGAGCGCTCATGTATGATGAGTGGTGGGACACATAAGCCA	660
Db	622	GAGGAAGTGCAGTTCGCGGCACTGCAGAGCGCTCATGTATGATGAGTGGTGGGACACATAAGCCA	681
OY	661	GAGAAAGCAGCCGCTGGGCCCTGTGGTGGTATGAGCCATATGAAGAACCCGACTTTTGGCCAC	720
Db	682	GAGAAAGCAGCCGCTGGGCCCTGTGGTGGTATGAGCCATATGAAGAACCCGACTTTTGGCCAC	741
OY	721	TTTCATGTATGAATAGTCTCTGGGAAACACAGACCTCTTCTTCATCCAGGCGCCAGGAG	780
Db	742	TTTCATGTATGAATAGTCTCTGGGAAACACAGACCTCTTCTTCATCCAGGCGCCAGGAG	801
OY	781	TACACCGTGGTGTTTTGGAGATGGAAGAAAGAGAGTCCAGGAATPACACGGTGTGAACACA	840
Db	802	TACACCGTGGTGTTTTGGAGATGGAAGAAAGAGAGTCCAGGAATPACACGGTGTGAACACA	861
OY	841	GAGAAAGCCCTCATGAGAGGTGCAGAGAGATGCTGCTCCCTGGGATGAAAGTGAAGCTCCAG	900
Db	862	GAGAAAGCCCTCATGAGAGGTGCAGAGAGATGCTGCTCCCTGGGATGAAAGTGAAGCTCCAG	921
OY	901	CTCCAGGTCCAGAAATCCCTGTGGACAGCCAGCAGAGACAGAGAAATCTPACATCTACAC	960
Db	922	CTCCAGGTCCAGAAATCCCTGTGGACAGCCAGCAGAGACAGAGAAATCTPACATCTACAC	981
OY	961	CTCAAAGGCGATGTGCCCCCTTAACACACCCCAACAGGCGTTGATATCTCCAGCTGTGCTC	1021
Db	982	CTCAAAGGCGATGTGCCCCCTTAACACACCCCAACAGGCGTTGATATCTCCAGCTGTGCTC	1041
OY	1021	ACCTGCTTCTTGGGCCGCTGCTTATTAAGAAATTCCTCACTGGTCTTACGGGGCTC	1081
Db	1042	ACCTGCTTCTTGGGCCGCTGCTTATTAAGAAATTCCTCACTGGTCTTACGGGGCTC	1101
OY	1081	GCCGATGGGGCTTGTGGCTGTATTTCCCGTGGTGGGGGACCCCAAGGACACTCTCC	1141
Db	1102	GCCGATGGGGCTTGTGGCTGTATTTCCCGTGGTGGGGGACCCCAAGGACACTCTCC	1161
OY	1141	TACCTGTGCTCACACACAGCCCAACAGGTCCAAGTTTCAGATCGCGGATGAAGACGACGG	1201
Db	1162	TACCTGTGCTCACACACAGCCCAACAGGTCCAAGTTTCAGATCGCGGATGAAGACGACGG	1221
OY	1201	CAGAAACCCCTACCCAGTGAAGAGCCATGGAAGTGGTCAACAGCGGCTCTGAAGTCTGGTAC	1261
Db	1222	CAGAAACCCCTACCCAGTGAAGAGCCATGGAAGTGGTCAACAGCGGCTCTGAAGTCTGGTAC	1281
OY	1261	AGCAATGGGCGCCGCGCTTCCTGTGATGTAGCTGCTCCCTGGAGATCTCAGAGCGGCTG	1321
Db	1282	AGCAATGGGCGCCGCGCTTCCTGTGATGTAGCTGCTCCCTGGAGATCTCAGAGCGGCTG	1341
OY	1321	GAGCCCTACATAGGCCCTCTCCATGTATACGTCAAGTCTGTGACAGCTCTGAAGGCGACAGGG	1381
Db	1342	GAGCCCTACATAGGCCCTCTCCATGTATACGTCAAGTCTGTGACAGCTCTGAAGGCGACAGGG	1401
OY	1381	GAGGAGGTCTGTGTGGCTTGGATGATGAAGCCCAACTCTTGGTATGTACCACTCTACAC	1441
Db	1402	GAGGAGGTCTGTGTGGCTTGGATGATGAAGCCCAACTCTTGGTATGTACCACTCTACAC	1461

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 01:49:41 ; Search time 3953 Seconds

(Without alignments)
16079.061 Million cell updates/sec

Title: us-09-836-392-8_COPY_22_2205

Perfect score: 2184

Sequence: 1 atctgcggcgcactgcgggc.....ctgcactgcagaagaag 2184

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database :

GenBank:*

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_ov:*

7: gb_ov:*

8: gb_ov:*

9: gb_ov:*

10: gb_ov:*

11: gb_ov:*

12: gb_ov:*

13: gb_ov:*

14: gb_ov:*

15: gb_ov:*

16: gb_ov:*

17: gb_ov:*

18: gb_ov:*

19: gb_ov:*

20: gb_ov:*

21: gb_ov:*

22: gb_ov:*

23: gb_ov:*

24: gb_ov:*

25: gb_ov:*

26: gb_ov:*

27: gb_ov:*

28: gb_ov:*

29: gb_ov:*

30: gb_ov:*

31: gb_ov:*

32: gb_ov:*

33: gb_ov:*

34: gb_ov:*

35: gb_ov:*

36: gb_ov:*

37: gb_ov:*

38: gb_ov:*

39: gb_ov:*

40: gb_ov:*

41: gb_ov:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2184	100.0	5370	9 AB058693	AB058693 Homo sapi
2	2176	99.6	4210	9 AK058028	AK058028 Homo sapi
3	1936.8	88.7	6045	6 AX250161	AX250161 Sequence
4	1915.4	87.7	3509	9 BC005408	BC005408 Homo sapi
5	1840	84.2	3112	9 AK054662	AK054662 Homo sapi
6	1129.8	51.7	2790	9 AK094556	AK094556 Homo sapi
7	858.8	39.3	110000	2 AL390202.01	Continuation (2 of
8	858.8	39.3	110000	2 AL390202.02	Continuation (3 of
9	858.8	39.3	178876	2 AC019254	AC019254 Homo sapi
10	850.6	38.9	183401	2 AC020578	AC020578 Homo sapi
11	850.6	38.9	1980	10 BC027119	BC027119 Mus muscu
12	775.8	35.5	2047	9 AK026772	AK026772 Homo sapi
13	611	28.0	288428	2 AC127595	AC127595 Mus muscu
14	477.6	21.9	526	6 AR175612	AR175612 Sequence
15	283	13.0	178529	2 AC090907	AC090907 Homo sapi
16	237.2	10.9	179377	2 AC117156	AC117156 Rattus no
17	225.2	10.3	80468	2 AC096233	AC096233 Rattus no
18	160.2	7.3	179377	2 AC117156	AC117156 Rattus no
19	126.2	5.8	871	4 CR002196	CR002196 Rattus no
20	69.4	3.2	2344	4 RABPKC	RABPKC
21	68	3.1	125020	9 AF429315	AF429315 Homo sapi
22	67.8	3.1	1830	9 HSPKCG	HSPKCG
23	67.8	3.1	2191	6 AX035149	AX035149 Sequence
24	67.8	3.1	2506	8 AF345987	AF345987 Homo sapi
25	66	3.0	59843	8 AP000816	AP000816 Oryza sat
26	65.4	3.0	161266	8 AP001278	AP001278 Oryza sat
27	65.4	3.0	198057	2 AC073719	AC073719 Mus muscu
28	64.6	3.0	2167	6 A37236	A37236 Sequence 5
29	64.6	3.0	2168	4 BOVPRG	BOVPRG
30	63.4	2.9	3030	10 BC034074	BC034074 Mus muscu
31	60.4	2.8	110000	2 LMFLCHR3_18	Continuation (19 of
32	59.4	2.7	125020	9 AF429315	AF429315 Homo sapi
33	59	2.7	7195	1 AF159501	AF159501 Myxococcu
34	58.8	2.7	2363	10 BC004839	BC004839 Mus muscu
35	58.4	2.7	110000	2 LMFLCHR16_05	Continuation (6 of
36	58.4	2.7	110000	2 LMFLCHR32_07	Continuation (8 of
37	58.2	2.7	1353	8 GONCDPK	GONCDPK
38	58	2.7	42160	3 AC005804	AC005804 Leishmani
39	57.8	2.7	110500	2 AC121231	AC121231 Leishmani
40	57.8	2.6	148533	2 AP005449	AP005449 Oryza sat
41	57.2	2.6	2225	6 AX035150	AX035150 Sequence
42	57.2	2.6	2649	6 AX035148	AX035148 Sequence
43	57.2	2.6	2810	6 AT069425	AT069425 Drosophil
44	57.2	2.6	3025	6 E04371	E04371 cDNA sequen
45	57.2	2.6	3025	10 RNPCKG	X07287 Rat mRNA to

ALIGNMENTS

RESULT 1

AB058693 5370 bp mRNA linear PRI 05-JUN-2001

LOCUS AB058693

DEFINITION Homo sapiens mRNA for KIAA1790 protein, partial cds.

ACCESSION AB058693

VERSION AB058693.1 GI:14017796

KEYWORDS

SOURCE

ORGANISM

Homo sapiens brain cDNA to mRNA, clone:fh24104.

Homo sapiens

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

Nagase,T., Nakayama,M., Nakajima,D., Kikuno,R. and Ohara,O.

Prediction of the coding sequences of unidentified human genes. XV.

The complete sequences of 100 new cDNA clones from brain which code

for large proteins in vitro
 JOURNAL DNA Res. 8 (2), 85-95 (2001)
 MEDLINE 21245130
 REFERENCE 2 (bases 1 to 5370)
 AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute,
 Department of Human Gene Research, 1532-3, Yama, Kisarazu, Chiba
 292-0812, Japan (E-mail: cdna.hf@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,
 Fax: 81-438-52-3914)

FEATURES

Source

Location/Qualifiers

1..5370

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="fh24104"

/tissue_type="brain"

/note="vector: pBluescriptII SK plus"

1..5370

/gene="KIAA1790"

<1..4112

/gene="KIAA1790"

/note="Start codon is not identified."

/codon_start=3

/product="KIAA1790 protein"

/protein_id="BAB47419.1"

/db_xref="GI:1401797"

/translation="PROGSTYLELQTRAPVYHGEATIRTTKWELOPAGRAKY
 ESEFNVMDIGPAMATVNCQFETDKALVYVWNLAGEAVANLQFWLNIKAP
 NAVLVYVTHDLIEAFERVRIATLRYVALCRSPSGATGPDITFEHLISC
 KSLGOGRLHRLIFHTCSMKDVGSTIGCRALRLIPRSVLSQEPVLAQORRSR
 DDVOYLDROLEQVEOTPNNDIKDYEDLOAIFLITETGLLFPDISHLRNLYE
 DPVLSCLRIENIKSRVAKNGVRAEDLRLNLVGTGTQTEDEQFQFLAFEL
 ALPVANOSYLLPHLPSKPGLDTHGMHRAVNTQVREKVFVGVGQAPRIYDEK
 LAEDLQLEFKKNTKSKRNKVTIYSEFGRNCRNCFEVRNQTIIWOELVTFPD
 GYLSVESSDVMKKKSKGKMTVCQSEVDRSAAFIDHVNSLIDQEPVAFDGEESR
 GTPMEQVPCVCEYTAHOHTDESEKEDVOYQMECVATLEVERISCPRPDLP
 VPLOEIVPELMDPDPARPLFENKLESEDESGVLOGGSGVATVYARQGVAVK
 RFIHKRRNPANPADTMRRLRLRTDANKNSERROESMLHALQHCYALIGTSH
 PLCFALBELAPLSTLNTVLEENARDSSFLPGLHMLTOKIAYIAGGLALHKKNIIFCD
 LKSDNITLWSDLVKEHINIKLSDGISHSOFHEALGVETGPGVQAPRIYDEK
 VDMFSYGVNIELLSQORPALGHQLOIAKLSKIRVLPVLOGPEVQFRALQALMEC
 WDTPERKPLALSVYQSKDPTFATFMEILCGKOTAFSSGOGGYVFPDGEESR
 NYTVNTEKIMEYORMCCPGMKVSCOLOYRMTATBEDOKIYITLKGMCPLNTPO
 QALDTPAVTGFCLAVPIYKKNISLYLALDGLVAVPYVSGTRKDSYSICSTANR
 SKESIADBARONPYPVAMEVNSGSVWISNBPGLVIDCASLEICRLEPTWAPS
 MVTYVCSSEGEYVWCLDDKANSIVMHSYTYQICARFCGVPSPRLMFPVRL
 DTEPPASHTANPVPEDGSIDVYSIMSELGTOILIHOSLIDYCSMSYSPPR
 QAAKPSPLSPSSPASSSVPESTDCEDMDHTGAAADREHDLTPMDGETFSOHLQ
 AVKTLAVRDLIWPVRGSDVIGLEKDESEORVLAIVLAKRELTPGVLDAAVVA
 KDYVCTFENENTMCLAVWRGKAREFDITYSYELGRLEACTRRKR"

BASE COUNT 1226 a 1553 c 1502 g 1089 t
 ORIGIN

Query Match 100.0%; Score 2184; DB 9; Length 5370;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGGACCTGCGGGCCACCGATGCAAGAACTTCTCGAGTTCCGGCAGAG 60
 DB 1926 ATGTGAGGACCTGCGGGCCACCGATGCAAGAACTTCTCGAGTTCCGGCAGAG 1985
 QY 61 GCCAGATGCTGACGCGGTGACACACCCCTGCATGTCGGCTCATCGGATCAGCATC 120
 DB 1986 GCCAGATGCTGACGCGGTGACACACCCCTGCATGTCGGCTCATCGGATCAGCATC 2045
 QY 121 CACCGCTGCTGCGCGGTGAGGCTGCGGCGCTCAGAGCTCAACACGCTGCTGC 180
 DB 2046 CACCGCTGCTGCGCGGTGAGGCTGCGGCGCTCAGAGCTCAACACGCTGCTGC 2105
 QY 181 GAGAAAGCAGAGATCTCTTATACCCCTGGGACACATGCTCAACCCAAAATAGCC 240
 DB 2106 GAGAAAGCAGAGATCTCTTATACCCCTGGGACACATGCTCAACCCAAAATAGCC 2165

QY 241 TACCAGATGCTGCGGCGCTGCTGCTTACTGACAGAAACATCATCTTCTGTGACCTG 300
 DB 2166 TACCAGATGCTGCGGCGCTGCTGCTTACTGACAGAAACATCATCTTCTGTGACCTG 2225
 QY 301 AAGTCGCAACATCTTCTGCTGCTGCTTACTGACAGAAACATCATCTTCTGTGACCTG 360
 DB 2226 AAGTCGCAACATCTTCTGCTGCTGCTTACTGACAGAAACATCATCTTCTGTGACCTG 2285
 QY 361 TCTGATGAGGATTTGAGGACGATCATCTGATGAGGCGGCTGAGGCGGCGACT 420
 DB 2286 TCTGATGAGGATTTGAGGACGATCATCTGATGAGGCGGCTGAGGCGGCGACT 2345
 QY 421 CTTGCTACACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 DB 2346 CTTGCTACACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2405
 QY 481 TTTCCCTATGGAATGCTCTTACAGATGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 DB 2406 TTTCCCTATGGAATGCTCTTACAGATGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAG 2465
 QY 541 CACGAGCTCAGATTTGCCAAGAGCTGTCCAAGGAGCATCCCGGCTTCTGGGCGAGCGG 600
 DB 2466 CACGAGCTCAGATTTGCCAAGAGCTGTCCAAGGAGCATCCCGGCTTCTGGGCGAGCGG 2525
 QY 601 GAGGAGTGCAGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 DB 2526 GAGGAGTGCAGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2585
 QY 661 GAGGAGTGCAGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
 DB 2586 GAGGAGTGCAGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2645
 QY 721 TTTATGATGAGATGCTGCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 DB 2646 TTTATGATGAGATGCTGCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2705
 QY 781 TACACCGTGGTGTGGGAG 840
 DB 2706 TACACCGTGGTGTGGGAG 2765
 QY 841 GAGAGGCGCTCAG 900
 DB 2766 GAGAGGCGCTCAG 2825
 QY 901 CTCCAGTTCAGAGATCCCTGTGGACACGACCGACGACGACGACGACGACGACGACGAC 960
 DB 2826 CTCCAGTTCAGAGATCCCTGTGGACACGACCGACGACGACGACGACGACGACGACGAC 2885
 QY 961 CTCAAGGCGATGCGCCCTTAAACACACGACCGACGACGACGACGACGACGACGACGAC 1020
 DB 2886 CTCAAGGCGATGCGCCCTTAAACACACGACCGACGACGACGACGACGACGACGACGAC 2945
 QY 1021 ACCTGCTTGTGGCGGCTGCTGTATTAAGAAATCTTCTACTGCTTGTAGCGGCGCTC 1080
 DB 2946 ACCTGCTTGTGGCGGCTGCTGTATTAAGAAATCTTCTACTGCTTGTAGCGGCGCTC 3005
 QY 1081 GCCGATGAGCTGTGCTGTGCTTCCGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 DB 3006 GCCGATGAGCTGTGCTGTGCTTCCGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3065
 QY 1141 TACCTGTCTACACACAGGCAACAGGCTCAAGTTCACATCGGAGTGAAGAGCGACAGG 1200
 DB 3066 TACCTGTCTACACACAGGCAACAGGCTCAAGTTCACATCGGAGTGAAGAGCGACAGG 3125
 QY 1201 CAGAACCCCTACAGTGAAGGCAATGAGAGTGTCAACAGGCGCTGAGGCTGTGTAC 1260
 DB 3126 CAGAACCCCTACAGTGAAGGCAATGAGAGTGTCAACAGGCGCTGAGGCTGTGTAC 3185
 QY 1261 AGCAATGGGCGGCGGCTCTGTCTATGATGATGATGATGATGATGATGATGATGATGATG 1320
 DB 3186 AGCAATGGGCGGCGGCTCTGTCTATGATGATGATGATGATGATGATGATGATGATGATG 3245

QY	1652	ACGTGCTGCATCTCTCTCTACCTCCATCCATCCCAACCCCGCAGAGGCGGACGATCCCCCTCA	1711
Db	102094	ACTGCTGCATCTCTCTCTACTCTTCATCTCCCAACCCCGCAGAGGCTCCACGCTCCCTCCAA	1020355
QY	1712	GCCTTCCCAAGTCCCCAGCAAGTTCTTCCAGTGTGCTTTCTCCACCGACTGCGAGACT	1771
Db	102034	GCCTTCCCAAGTCCCCAGCAAGTCTTCCAGTGTGCTTTCTCCACCGACTGCGAGAGACT	1019795
QY	1772	CAGACATGCTACTATACACC CGCGTCTGCTGCTCTCCGACGAGTGTGAGCATGATGACTGACCCCA	1831-
Db	101974	CAGACATGCTACTATACAGCGCGGTCTGCTGCTCTCCGACGAGTGTGAGCATGATGACTGACCCCA	101915
QY	1832	TGGACGGGGAGACCTTACAGCCAGACCTGTCAGAGCGCGTGAAGATCTGCGCGTCAAGAAC	1891
Db	101914	TGGACGGGGAGACCTTACAGCCAGACCTGTCAGAGCGCGTGAAGATCTGCGCGTCAAGAAC	101855
QY	1892	TCATTTGGGTCCCGCAGCGCGCGTGGAGATGTTATCTCATTTGGCGCTGGAAGAAGATCTGT	1951
Db	101854	TCATTTGGGTCCCGCAGGATGTTTCCGAGAGTGAGGGCACACATCCAGAGGCGACGCGCACTG	101795
QY	1952	AAGCCACAGCGCGCGCGAG	1969
Db	101794	CTCTGCTCTGGGACAG	101777

RESULT - 8
AL390202_02/c
WPCOMMENT

Sequence split into 10 fragments LOCUS AL390202 Accession AL390202

Fragment Name	Begin	End
AL390202_00	1	110000
AL390202_01	100001	210000
AL390202_02	200001	310000
AL390202_03	300001	410000
AL390202_04	400001	510000
AL390202_05	500001	610000
AL390202_06	600001	710000
AL390202_07	700001	810000
AL390202_08	800001	910000
AL390202_09	900001	988176

Continuation (3 of 10) of AL390202 from base 200001 (AL390202 Homo sapiens chromosome 2)

Query Match	39.3%	Score 858.8;	DB 2;	Length 110000;
Best Local Similarity	96.0%	Pred. No. 1.6e-145;		
Matches 881; Conservative	0;	Mismatches 37;	Indels 0;	Gaps 0;

QY	1052	AGAAATTCCTACCTGGTCTTACGCGGCTCGCGGATAGGAGCTTGAGCTGTGTTCCCGTGG	1111
QY	1112	TGCGGGGACCCCAAGAGACAGCTGCTCTTACCTGTGCTCACACACAGCAACAGGTCCA	1171
QY	1172	AGTTCAGATGGCGGAGTAAACGACGCGCAGAACCCCTACCACAGTGAAGCCATGGAGG	1231
QY	1232	TGGTCACACAGCGGCTCTGAGGCTGCTGACAGCAATGGGCGCGGCTCCTTGTCATCGACT	1291
QY	1292	TGGTCACACAGCGGCTCTGAGGCTGCTGACAGCAATGGGCGCGGCTCCTTGTCATCGACT	1351
QY	1352	CAGTGTGTGACAGCTCTGAGGCGCAGAGGGGAGAGAGTGTGTGTGCTGTGATGACAAAG	1411
QY	1352	CAGTGTGTGACAGCTCTGAGGCGCAGAGGGGAGAGAGTGTGTGTGCTGTGATGACAAAG	1411
QY	1412	CCAACTCCTTGGTATATGACACTCCACCACCTACACACTGTGTGCCCGGTACTTCTGCG	1471

Db	2334	CCAACCTCCTTGGTGTGTGACACATCCACACCTACACACTGTGATCCCGGACTCTCTCG	2275
QY	1412	GGGTCGCCAGGCCCTCTCAGGGAACAATGTTCCCGCGGCCCTTGACACAGGAACCCCGG	1531
Db	2274	GGGTCCCAAGCCCCCTCCAGGGAACATGTTCCCGTGGCCCTTGGAACAGGAACCCCGG	2215
QY	1532	CAGCCAGCCACACGCGCCACCAACCMAAGGTGCTGAGGGGGACTCCATCCGCGACGTGAGCA	1591
Db	2214	CAGCCAGCCACACGCGCCAAACCAAGGTGCTGAGGGGGACTCCATCCGCGACGTGAGCA	2155
QY	1592	TCATGTACAGTGAAGAGAGCTGGGACAGCAGATTCGTATCCACAGGAATCACTACTGACT	1651
Db	2154	TCATGTACAGTGAAGAGAGCTGGGACAGCAGATTCGTATCCACAGGAATCACTACTGACT	2095
QY	1652	ACTGCTCCATGTCTCTCTACACTCTCATATCCCAACCCCGCAGGCTCCGAGTCCCCCTTAA	1711
Db	2054	ACTGCTCATGTCTCTCTCTACACTCTCATATCCCAACCCCGCAGGCTCCGAGTCCCCCTTAA	2035
QY	1712	GGCTCCCGCAGCTCCCGACAGCAAGTTCTTCCATGTGCTCTTCTCCACAGCATGCGAGACT	1771
Db	2034	GGCTCCCGCAGCTCCCGACAGCAAGTTCTTCCATGTGCTCTTCTCCACAGCATGCGAGACT	1975
QY	1772	CAGACATGCTCATACAGCGCCGCTGCTGCTCCGACGAGTCTGAGGATACCTGACCCCA	1831
Db	1974	CAGACATGCTCATACAGCGCCGCTGCTGCTCCGACGAGTCTGAGGATACCTGACCCCA	1915
QY	1832	TGGACGGGGGAACCTTTCAGCCAGACACTGCAAGGCGTAAAGATCCTGCCCGTCAGAGACC	1891
Db	1914	TGGACGGGGGAACCTTTCAGCCAGACACTGCAAGGCGTAAAGATCCTGCCCGTCAGAGACC	1855
QY	1892	TCATTTGGGTCCCGCAGCGCGGTGGAGATGTTATTCGTATTGGCCTGGAGAGAAGATTTCG	1951
Db	1854	TCATTTGGGTCCCGCAGCGGTGGAGATGTTATTCGTATTGGCCTGGAGAGAAGATTTCG	1795
QY	1952	AAGCCACGCGGGCGCCGAG	1969
Db	1794	CTCCGTGCTGTGGGGACAG	1777

RESULT	9
ACOL19254	
LOCUS	ACOL19254
DEFINITION	Homo sapiens chromosome 15 clone RP11-424119 map 15, *** SEQUENCING IN PROGRESS ***; 3 unoriented pieces.
ACCESSION	ACOL19254
VERSION	ACOL19254.9 GI:22123259
KEYWORDS	HNG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 178876) Blatten,B., Nusbaum,C. and Lander,E. Homo sapiens chromosome 15, clone RP11-424119 Unpublished 2 (bases 1 to 178876)
REFERENCE	Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boguslavsky,I., Boukhalter,B., Brown,A., Burnett,G., Castle,A., Chepel,X., Colangelo,M., Collins,S., Collamore,A., Cooke,P., DeArillano,K., Dewar,K., Domino,M., Doyle,M., Fenster,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J., Gardy,R.S., Grant,G., Hagos,B., Heatford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kahn,L., Karatas,A., Kiehn,J., Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K., McPheters,R., Meldrim,J., Menenius,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Plierre,N., Plisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy A., Santos,R., Severy,P., Spencer,B., Stange-Rothman,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
AUTHORS	

TITLE	Direct Submission
JOURNAL	Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 178876)
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Birnbaumer,M., Brown,D., Chao,R., Church,G.M., Cook,W.R., DeRubeis,H., Dorschner,J.A., Ewing,L., Fulton,T.S., Gage,F.H., Gill,R., Glavinskas,I., Gruber,T.P., Harlowe,B., Heerdt,J.G., Hillier,L.G., Holt,R.A., Hughes,S.D., Hunkeler,K., Jackson,A.R., Jaiswal,S., Jones,R.C., Kalishy,J.M., Kaur,V., Kennedy,K.L., Kim,Y.J., Kirkpatrick,M.O., Kohane,D.E., Krangel,M.B., Kravits,S., Li,X., Lin,S.M., Liu,X., McEwen,G.F., McPherson,R., Miller,R., Muzny,D.B., Nelson,D.L., Nussenzweig,V., O'Connell,K., Park,S., Petrij,R., Peterson,J., Phillips,R.D., Pohlman,R., Quake,C.R., Reichert,H., Roberts,R., Roeder,B., Sanger,F., Schaefer,A., Scherer,P.E., Sherry,B.T., Skusek,D., Smith,R.D., Stange,C., Steinberg,M., Stewart,R., Thomsen,M., Tringali,G., Upton,K., Venter,A., Wang,J., Wang,M., Wang,Z., Weinstock,M., Westphal,K., Whittaker,P., Winkler,C., Wu,M., Yan,H., Zannis-Hadjopoulos,D.

TITLE Direct Submission
JOURNAL Submitted (06-ANG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Aug 6, 2002 this sequence version replaced 91:21699484. All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

*	I	58973:	contlg of 58973	bp in length
*		58974	59073: gap of	100 bp
*		59074	91153: contlg of 32080	bp in length
*		91154	91253: gap of	100 bp
*		91254	178876: contlg of 87623	bp in length.

[illegible]

LOCUS	AC020578	183401 bp	DNA	linear	HTG 15-FEB-2001
RESULT 10					
AC020578/c					
DEFINITION	AC020578	183401 bp	DNA	linear	HTG 15-FEB-2001
	Homio sapiens chromosome 15 clone RP11-505E24, WORKING DRAFT				
ACCESSION	AC020578				
VERSION	AC020578.8	GI:18831370			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	Homio sapiens.				
ORGANISM	Homio sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 183401)				
AUTHORS	Abola,A.P., Bruno,D., Conn,L., Della Rosa,M., Faulkner,D.,				
	Federspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,				

JOURNAL
AUTHORS

Mao, J., Lam, B., Marathe, R., Miranda, M., Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D., Southwick, A.M., Wilhelmy, J., Yu, S. and Davis, R.W.

TITLE
JOURNAL

2 (bases 1 to 183401)
Submitted (05-JAN-2000) DNA Sequencing and Technology Center,
Stanford University, 835 California Avenue, Palo Alto, CA 94304,
USA

COMMENT

On Feb 15, 2001 this sequence version replaced g1:12830184.

Center: Stanford DNA Sequencing and Technology Development

Center code: SDBTDC

Web site: <http://sequence-www.stanford.edu/group/human/>

Contact: hum-info@sequence.stanford.edu

Project Information

Center project name: 754

Center clone name: RP11-505E24

Sequencing Vector: M13mp18; X02513

Chemistry: Dye-terminator; 0% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 178788 bases at least Q40

Consensus quality: 180438 bases at least Q30

Consensus quality: 181007 bases at least Q20

Insert size: 174146; agarose-fp

Insert size: 183101; sum-of-contigs

Quality coverage: 9.8x in Q20 bases; agarose-fp

Quality coverage: 9.3x in Q20 bases; sum-of-contigs.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 3433: contig of 3433 bp in length

* 3434 3533: gap of unknown length

* 3534 16557: contig of 13124 bp in length

* 16558 16757: gap of unknown length

* 16758 28859: contig of 12102 bp in length

* 28860 28959: gap of unknown length

* 28960 183401: contig of 154442 bp in length.

* Location/Qualifiers

1. 183401

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="15"

/clone="RP11-505E24"

/clone_1lb="RP11 human BAC library 11"

1. 3433

/note="assembly_name:Contig17"

3534. 16557

/note="assembly_name:Contig19"

16758. 28859

/note="assembly_name:Contig20"

clone_end:17"

28960. 183401

/note="assembly_name:Contig21"

clone_end:SP6"

BASE COUNT 49346 a 42604 c 44333 g 46782 t 346 others

ORIGIN

Query Match 39.3%; Score 858.8; DB 2; Length 183401;
Best Local Similarity 96.0%; Pred. No. 1.5e-149;
Matches 881; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

1052 AGAATTCCTACCTGGTCTTAGGGGCGCTCCGATGGGCTTGCTGCTCCGCTGG 1111
|||||
Db 60707 AGAATTCCTACCTGGTCTTAGGGGCGCTCCGATGGGCTTGCTGCTCCGCTGG 60648
1112 TCGGGGGACCCCAAGAGACAGCTCTCTCTACCTGTCTCACACAGCAAGCTCA 1171
|||||
Db 60647 TCGGGGGACCCCAAGAGAGACAGCTCTCTCTACCTGTCTCACACAGCAAGCTCA 60588
1172 AGTTAGCATTCGGGATGAGAGACAGGACAGGACAGGACAGGACAGGACAGG 1231
|||||
Db 60587 AGTTAGCATTCGGGATGAGAGACAGGACAGGACAGGACAGGACAGGACAGG 60528
1232 TGGTCAACAGCGGCTCTGTGAGCTGTGATACAGCAATGGGCGGCTCTGTGATGACT 1291
|||||
Db 60527 TGGTCAACAGCGGCTCTGTGAGCTGTGATACAGCAATGGGCGGCTCTGTGATGACT 60468
1292 GTGCTCTCTGAGATCTGACAGGCGGCTGAGACCTTACATGGCCCTCCATGGTTAGCT 1351
|||||
Db 60467 GTGCTCTCTGAGATCTGACAGGCGGCTGAGACCTTACATGGCCCTCCATGGTTAGCT 60408
1352 CAGTGTGTGACGCTCTGAGGCGCAGAGGAGAGAGAGTCTGTGTGCTGTGATGACAGG 1411
|||||
Db 60407 CAGTGTGTGACGCTCTGAGGCGCAGAGGAGAGAGTCTGTGTGCTGTGATGACAGG 60348
1412 CCAACTCTTGTGTGATGTACACATCCACACCTTACAGCTGTGTGCTGTGCTGTGCTGG 1471
|||||
Db 60347 CCAACTCTTGTGTGATGTACACATCCACACCTTACAGCTGTGTGCTGTGCTGTGCTGG 60288
1472 GGGTCCCGACGCCCCCTCAGGAGCATGTTTCCGCTGGCCCTTGGACAGGACCCCGG 1531
|||||
Db 60287 GGGTCCCGACGCCCCCTCAGGAGCATGTTTCCGCTGGCCCTTGGACAGGACCCCGG 60228
1532 CAGCCAGCCACAGGCGCAACCCAAAGGTGCTGAGGGGAGCTCCATCGGGAGGTGAGCA 1591
|||||
Db 60227 CAGCCAGCCACAGGCGCAACCCAAAGGTGCTGAGGGGAGCTCCATCGGGAGGTGAGCA 60168
1592 TCATGTACAGTGAAGAGCTGGGCGAGCAGATCTGTATCCACAGGATCTACTGACT 1651
|||||
Db 60167 TCATGTACAGTGAAGAGCTGGGCGAGCAGATCTGTATCCACAGGATCTACTGACT 60108
1652 ACTGCTCATGTCTCTCTACTCTCTATCCACCCAGCCGAGGCTGCCAGTCCCTCAA 1711
|||||
Db 60107 ACTGCTCATGTCTCTCTACTCTCTATCCACCCAGCCGAGGCTGCCAGTCCCTCAA 60048
1712 GCTCTCCAGCTCCCGACAGAAATGCTTCCAGTGTGCTTCCACAGGCTGGAGGACT 1771
|||||
Db 60047 GCTCTCCAGCTCCCGACAGAAATGCTTCCAGTGTGCTTCCACAGGCTGGAGGACT 59988
1772 CAGACATGCTACATACGCGCGGTGCTGCTCCGACAGGCTGTGAGCTGACCTGACCCCA 1831
|||||
Db 59987 CAGACATGCTACATACGCGCGGTGCTGCTCCGACAGGCTGTGAGCTGACCTGACCCCA 59928
1832 TCGAGCGGAGACCTTACAGCAGCAGACCTGAGAGCTGTGAGAGCTGCTGCTGACAGACC 1891
|||||
Db 59927 TCGAGCGGAGACCTTACAGCAGCAGACCTGAGAGCTGTGAGAGCTGCTGCTGACAGACC 59868
1892 TCATTGGGTCTCCCGAGCGCGGTGAGGATGTATCTCATTTGGCTTGGAGAGAGATTCTG 1951
|||||
Db 59867 TCATTGGGTCTCCCGAGCTGCTTCCGAGGTGAGGACACATCCAGGCGAGCCCATCTG 59808
1952 AAGCCAGCGGCGCGAG 1969
|||||
Db 59807 CTTCTGCTCTGGGACAG 59750

RESULT 11
BC027199 1980 bp mRNA linear ROD 07-AUG-2002
LOCUS BC027199
DEFINITION Mus musculus, similar to hypothetical protein FLJ23119, clone
ACCESSION BC027199
VERSION BC027199.1 GI:20073184

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT

MGC.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1980)
Strausberg, R.
Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amgbc@bcm.tmc.edu
Guanarane, P.H., Garcia, A.M., Lu, X., Huijy, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://lmage.llnl.gov>
Series: IRAX Plate: 37 Row: k Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1. 1980
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:28646 IMAGE:4224922"
/tissue="type="Colon, normal, 5 month old male mouse."
/clone_id="NCI CGAP_C024"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
395. 1153
/codon_start=1
/product="Similar to hypothetical protein FLJ23119"
/protein_id="AAH27119.1"
/db_xref="gi:20073185"
/translation="MHSATYQLCARFYCGDPNPLRDFSVOPSLVETPGSHKTTSG
PVEECIAVDSIMYSEELGTQLITFDSDLDICSMSSYSSPHODRSPSSLSLTS
YSSVPSANEDSDRLQEPSTSDRTEHLSPMOETFSOHLQAVKVAVDLWVPR
HGDIIVIGLEKDSGARGRVIAVKAELNHRGVLVAAVAAKDTVVCGRANETEM
CLAVWRGAREPFIIFYSELEGLRELECKTKRR"

CDS
BASE COUNT 517 a 531 c 535 g 397 t
ORIGIN

Query Match 38.9%; Score 850.6; DB 10; Length 1980;
Best Local Similarity 84.2%; Pred. No. 1e-147;
Matches 971; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

QY 1032 GGCCTGCTGTTATTAATAAATTCCTACCTGCTTACGAGGCGCTCGCGATGGCT 1091
DB 1 GGCAGTACTGTTATCAAAAAGAACCTCTCTGCTGCGAGGCTGACGAGCT 60
QY 1092 TGTGGCTGTTCCTCGGTGCGGGGACCCCAAGAGACAGTGTCTACCTGTGTC 1151
DB 61 CGTGGCTGTTCCTCTGTGCGACGGGGCACCCCAAGAGAGTGTCTACCTGTGTC 120
QY 1152 ACACACAGCAACAGGTCCAGTTCAGATCGGAGTGAAGACGACGAGAAACCCCTA 1211
DB 121 CCACACAGCAACAGGTCCAGTTCAGATCGGAGTGAAGACGAGAAACCCCTA 180
QY 1212 CCCAGTGAAGGCCATGAGAGTGTGTAACAGCGGCTCTGAGGTCTGTACAGCAATGGGCC 1271

DB 181 CCCAGTGAAGGCCATGAGAGTGTGTAACAGTGGCTCTAGCTCTGTACAGTAAAGCGCC 240
QY 1272 GGGCTCTCTTCATCATGCACTGCTGCTCCCTGAGATTCGACAGGCGCTGAGGCCCTACAT 1331
DB 241 TGGCTCTCTTCATCATGCACTGCTGCTCCCTGAGATTCGACAGGCGCTGAGGCCCTACAGC 300
QY 1332 GAGCCCTCTCATGTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1391
DB 301 AGCTTCGCTCATGTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 1392 CTGGTGGCTGATGACAGAGGCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
DB 361 CTGGTGGCTGATGACAGAGGCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 1452 GTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
DB 421 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 1512 CTGGAACAGGAAACCCCGGACGCCACAGCCCAACCCCAAGAGTGTCTGAGGGGGA 1571
DB 481 TTCACTCTAGAAACCC--CAGCAGCTCACAAACACCTCAAGAGGGGCTGTGGAAGA 537
QY 1572 CTCGATCGCGAGCGAGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1631
DB 538 GTGATCGCGAGCGAGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
QY 1632 CCAGAAATCAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1691
DB 598 CCAGAAATCAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
QY 1692 GGGTCCAGAGTCCCGCTCAAGCCCTCCAGCTCCCGAGCAAGTCTTCCAGTGTGCTTT 1751
DB 658 AGATCTGAGTCCCGCTCCAGCTCCCGAGCTCCCGAGCAAGTCTTCCAGTGTGCTTT 717
QY 1752 CTCGATCGCGAGCGAGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1811
DB 718 CTCGATCGCGAGCGAGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
QY 1812 TGAGCATGCTGAGAGCCCGCTGAGAGGAGAGCTTCAAGCCAGCACTGCAAGCCGAGAA 1871
DB 778 TGAGCATGCTGAGAGCCCGCTGAGAGGAGAGCTTCAAGCCAGCACTGCAAGCCGAGAA 837
QY 1872 GATCTCGCGCTGAGAGACCTCATTTGGGTCCCAAGCGCGGCTGAGAGATGTTATGTCAT 1931
DB 838 GGTCTGCTGCTCAAGAGACCTCATTTGGGTCCCAAGCGCGGCTGAGAGATGTTATGTCAT 897
QY 1932 TGGCTGAGAGAGATTTGAGAGCCCGAGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1991
DB 898 CGGCTGAGAGAGATTTGAGAGCCCGAGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 957
QY 1992 AGAGTGAATCCGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2051
DB 958 AGAGTGAATCCGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
QY 2052 GTGCACTTTGAAATGAAGAACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2111
DB 1018 GTGCACTTTGAAATGAAGAACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
QY 2112 CAGGAGTTCGACATTTCTACAGTCTGAGAGAGAGTGGGCGGCTGAGAGCTGTGAC 2171
DB 1078 CAGGAGTTCGACATTTCTACAGTCTGAGAGAGAGTGGGCGGCTGAGAGCTGTGAC 1137
QY 2172 TCGCAAGAGAGG 2184
DB 1138 TCGCAAGAGAGG 1150

RESULT 12
AK026772 2047 bp mRNA linear PRI 29-SEP-2000
LOCUS AK026772 Homo sapiens CDNA: FLJ23119 fls, clone LMG07978.
DEFINITION AK026772
ACCESSION AK026772

D		241	CTGACTACTGTCCATCATGCCTCCTACTACTCTCATCCCAACCCCGGCCAGAGCTCACAGTCCC	300
OY	1706	CCTTAAGCCTCCCCAAGCCTCCCAAGAATACTTCGACAAGTGCCCTTTCCCAACCAGCTGG	1765	
D	301	CCTAAAGCTCTCCCAAGCTCCCAAGAAATCTTTCAGTAGTGCTCTTCTCCACCACTGG	360	
OY	1766	AGGACTCAGACATGCTACATACAGCCCGGTGCTGCCCTCCAGAGGTCTGAGCATYACCTGA	1825	
D	361	AGGACTCAGACATGCTACATACAGCCCGGTGCTGCCCTCCAGAGGTCTGAGCATYACCTGA	420	
OY	1826	CCCCCATGAGAGGGAGAACCTTTCAGCCAGACACTCTCAGGCCGTGAAAATCTTCGGCGTCA	1885	
D	421	CCCCCATGAGAGGGAGAACCTTTCAGCCAGACACTCTCAGGCCGTGAAAATCTTCGGCGTCA	480	
OY	1886	GAGACCTATTGTGGGTGCCCAAGGCCGGGTGAGATGTTATCTCATATTGGCTCGAGAAAG	1945	
D	481	GAGACCTATTGTGGGTGCCCAAGGCCGGGTGAGATGTTATCTCATATTGGCTCGAGAAAG	540	
OY	1946	AATTGTGAAGCCAGAGGGGGCGAGTCAATTGCCGTCTTAAAAAGCCCGAGAGCTGACTCCG	2005	
D	541	AATTCAGGGCCAGAGGGGGCGAGTCAATTGCCGTCTTAAAAAGCCCGAGAGCTGACTCCG	600	
OY	2006	ATGGGCTCTGTTGGATCTGCTGCGGGGTGGCAAAGACACTGTGGTGTGACCTTTGAAA	2065	
D	601	ATGGGCTCTGTTGGATCTGCTGCGGGGTGGCAAAGACACTGTGTGTGTGACCTTTGAAA	660	
OY	2066	ATGAAACACAGAGTGTGCTGCTGGGCGCTGTGAGAGGGGCTGGGGCGCCAGAGAGTTGACA	2125	
D	661	ATGAAACACAGAGTGTGCTGCTGGGCGCTGTGAGAGGGGCTGGGGCGCCAGAGAGTTGACA	720	
OY	2126	TTTTCTACCACTCCTCAGAGAGCTGGGGCGGCTGTGAGGCTTGACACTGCCAAGAGAGG	2184	
D	721	TTTTCTACCACTCCTCAGAGAGCTGGGGCGGCTGTGAGGCTTGACACTGCCAAGAGAGG	779	
RESULT_13				
LOCUS	AC127595	288428 bp	DNA linear HTGS_11-AUG-2002	
DEFINITION	Mus musculus chromosome UNK clone RP23-49I1B, WORKING DRAFT			
ACCESSION	AC127595			
VERSION	AC127595.2 GI:22203893			
KEYWORDS	HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLPROP.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	McPherson,J.D. and Waterston,R.H.			
JOURNAL	The sequence of Mus musculus clone			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 288428)			
TITLE	McPherson,J.D. and Waterston,R.H.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (11-JUL-2002) Genome Sequencing Center, 4444 Forest Park			
AUTHORS	Parkway, St. Louis, MO 63108, USA			
TITLE	3 (bases 1 to 288428)			
JOURNAL	McPherson,J.D. and Waterston,R.H.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (11-AUG-2002) Genome Sequencing Center, 4444 Forest Park			
TITLE	Parkway, St. Louis, MO 63108, USA			
JOURNAL	On Aug 11, 2002 this sequence version replaced gi:21887033.			
COMMENT				
	Genome Center			
	Center: Washington University Genome Sequencing Center			
	Center code: WUGSC			
	Web site:http://genome.wustl.edu/gsc/index.shtml			
	Contact: submissions@watsn.wustl.edu			
	Project Information			
	Center project name: M.BA0049118			
	Summary Statistics			
	Sequencing vector: M13; 0%			

Best Local Similarity	82.8%	Pred	No. 1,2e-103	
Matches	711: Conservative	0: Mismatches	145: Indels	3: Gaps
QY 1050	AAAGAAATTCCTACCTGTGTGCTTTAGCGGGCCCTGGCCGATGGGCTTTGTGCTGTTCCTT	1109		
Db 190466	ACAGAACTCCTTCCTGGTGTGTGAGAGGCTGGCTGACAGCACTCGTGTGTTCCTGT	190525		
QY 1110	GGTGGGGGGGACCCCAAGGACAGCTGCTCTTACTGTGTCTACACACAGCCACAGCTC	1169		
Db 190526	GGCAGGGGGGACCCCAAGGAAAGCTGCTCTTACTGTGTCTCCACACAGCCACAGCTC	190585		
QY 1170	CAAGTTCAGCATTCGGGTGTAAGAGCGACGGCAGAACCCCTTACCAGTAAGAGGCATGGA	1229		
Db 190586	CAAGTTCAGCATTCGGGTGTAAGAGCGACGGCAGAACCCCTTACCAGTAAGAGGCATGGA	190645		
QY 1230	GGTGGTCAACAGCGGCTCTGAGGTCTGTGTACAGCAATGGCCGGGCTCTTGTTCATGGA	1289		
Db 190646	AGTGGTCAACAGGTGGGTGTGAGGTCTGTGTACAGCAATGGCCGGGCTCTTGTTCATGGA	190705		
QY 1290	CTGTGCTCCTCCCTGGAGATCTGACAGGCGGCTGAGCCCTTACATAGGCCCTCCATGGTTAC	1349		
Db 190706	CTGACACCATCTCTGGACATACACAGCGGCTGAGGCCCTTACACAGCTTCGTCATGGTTAC	190765		
QY 1350	GTCAGTGGTGGAGCTCTGAGGGGACAGGGGAGGAGGTGCTGTGGTCTCGATGAGCAA	1409		
Db 190766	GTCAGTGGTGGAGCTCTGAGGGGACAGGGGAGGAGGTGCTGTGGTCTCGATGAGCAA	190825		
QY 1410	GGCCAACTCTTGTGGATGATACACTCCACACATTCACAGCTGTGTGGCCGGTACTTGTG	1469		
Db 190826	GGCCAACTGTCTTATGATGATACACTCCACACATTCACAGCTGTGTGGCCGGTACTTGTG	190885		
QY 1470	CGGGGTCCCCAGCCCCCTCAGGGAGCATGTTCCGTGCGGCCCTTGGACACAGGAACCCC	1529		
Db 190886	TGGGGACCCCAATCCTCTCAGGGACACAGTTTCTGTGAGCCTTCAGTCTTGAAGAACCC	190944		
QY 1530	GGCAGCCAGCCACACAGGCCAACCCAAAGTGGCTTGAAGGGGAGACTCCATTCGGGAGGTGAG	1589		
Db 190945	--CAGCGAGTCACAAACACCACCTCAAAAGGGGCTGTGAAAGAGTCAATCGCGGATGTGAG	191002		
QY 1590	CATCATGTACAGTGTAGAGAGTGGGCGACCCAGATCCTGTATCCACACAGATCACTCATGA	1649		
Db 191003	CATCATGTACAGAGCGAGAGTGGGCGACCGAAGATCTCTACCCACACAGAGATGCTCACGGA	191062		
QY 1650	CTACTGTCTCATGTCTCTCTTACTCTCTCATATCCCAACCCCGCACAGGCTGCGAGCTCCCTTC	1709		
Db 191063	CTACTGTCTCATGTCTCTTATTTCTTGTGTCGCCACCCACACAGATTCCTAGTCTCCCTTC	191122		
QY 1710	AAGCTCCCGACGCTCCCGACGAAAGTCTTCCAGTGTGGCTTCTCCACCGACTGTGGAGA	1769		
Db 191123	CAGTCTGCCACGCTCCCTCACAGCTATTTCCAGTGTGCTTCTCTGCCACATATGAGAA	191182		
QY 1770	CTCAGACATGCTTACATAGACCCCGGTGTCCTCCGACAGAGTCTGTGACATGACTGACCC	1829		
Db 191183	CTCAGACAGGCTGTGAGAGCCCGAGTGTACCTCTGACAGAACTGTAGCATGACTGAGCC	191242		
QY 1830	CATGACAGGGGAGACCTTACAGCCACGACTGTGAGCGCTGTGAAGATCTTGGCGGTAGAGA	1889		
Db 191243	CATGATATGGGGAGAGTTCATGCGCACACCTGTGAGAGGTGCTTGTGCTTCAAGAA	191302		
QY 1890	CCTCATTTGGGTCGCCAGG 1908			
Db 191303	CCTCATTTGGGTCCTTAgg 191321			
RESULT 14				
LOCUS	AR175612	526 bp	DNA	linear
DEFINITION	Sequence 10 from patent US 6309849.			pat 17-DEC-2001
ACCESSION	AR175612			
VERSION	AR175612.1	GI:17916911		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			


```
*      89662 89761: gap of 100 bp
*      89762 133081: contig of 43320 bp in length
*      133082 133181: gap of 100 bp
*      133182 161268: contig of 28087 bp in length
*      161269 161368: gap of 100 bp
*      161369 165097: contig of 3729 bp in length
*      165098 165197: gap of 100 bp
*      165198 177624: contig of 12427 bp in length
*      177625 177724: gap of 100 bp
*      177725 178529: contig of 805 bp in length.
```

FEATURES
source 1. 178529
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RPCI-II Human Male BAC"

BASE COUNT 48864 a 42402 c 41950 g 44500 t 813 others
ORIGIN

Query Match 13.0%: Score 283; DB 2; Length 178529;
Best Local Similarity 100.0%; Pred. No. 7.6e-43;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 192 AGATCTTCCTTATACCCCTGGACACATGCTCACCCAAAATAAGCTACAGATCGC 251
|||||
Db 18911 AGATCTTCCTTATACCCCTGGACACATGCTCACCCAAAATAAGCTACAGATCGC 18852

OY 252 CTGGGGCCTGGCCCTACCTGACACAGAAAAACATCTTCTGTGACCTGAGTCGACAA 311
|||||
Db 18851 CTGGGGCCTGGCCCTACCTGACACAGAAAAACATCTTCTGTGACCTGAGTCGACAA 18792

OY 312 CATTCGTGTGTGTCCTGAGTCAAGAGACATCAACATCAAGCTCTGACTAGCG 371
|||||
Db 18791 CATTCGTGTGTGTCCTGAGTCAAGAGACATCAACATCAAGCTCTGACTAGCG 18732

OY 372 GATTTCGAGCAGTATTCATGAGGGGCCCTAGGCGCTGAGGGGACCTCTGCTACCA 431
|||||
Db 18731 GATTTCGAGCAGTATTCATGAGGGGCCCTAGGCGCTGAGGGGACCTCTGCTACCA 18672

OY 432 GGGCCGAGATCAGGCGCTGCATTTGTATATGATGAGAGTA 474
|||||
Db 18671 GGGCCGAGATCAGGCGCTGCATTTGTATATGATGAGAGTA 18629
```

Search completed: April 15, 2003, 05:21:35
Job time : 5379 secs

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 859 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-395-580-2

Alignment Scores:

Score:	1,54e-14	Length:	859
Percent Similarity:	295.50	Matches:	168
Best Local Similarity:	34.278	Conservative:	89
Query Match:	22.408	Mismatches:	257
	7.28%	Indels:	236
		Gaps:	31

US-09-836-392-8_COPY_22_2205 (1-2184) x US-08-395-580-2 (1-859)

QY 28 GCATGAAAGACTTCCTCCGAGTCCGGCAG---GAGGCCAGATGCTGACGCGCTGCAG 84
 DB 150 AAlaValIysLysValAlaArgAspLeuLysGluThrAspIleLysHisLeuArgLysLeuLys 169
 QY 85 CACCCCTGCATCGGGCGCTCATCGCATCAGATCCAGATCCACCGGCTGCTGCGC----- 138
 DB 170 HisProAsnIleIleThrPheLysGlyValCysThrGlnAlaProCysTyrCysIleLeu 189
 QY 139 CTGAGACTCGCGCGCTCAGACAGCTTACACCGTGTCTCCGAGAAGCGCAGATTCT 198
 DB 190 MetGluPheCysAlaGlnGlyGlnIleuThrGluValLeuArgAlaGly----- 206
 QY 199 TCTTTATACCTTCGGGACACATGCTTACCCAAAATAAGCTTACAGATGCGCTCGGC 258
 DB 207 -----ProValThrProSerLeuLeuValAspTrpSerMetGlyIleAlaGlyGly 223
 QY 259 CTGAGCTACCTGCAGCAAGAAAACATCATCTCTCTGACCTGAGTGGACAACTTCG 318
 DB 224 MetAsnTyrIleuHisLeuHisLysIleIleHisArgSerPheLysSerProAsnMetLeu 243
 QY 319 GTGTGCTCCCTGACGTCAAGAGACATCAATCAAGTATGATCTGACGCGATTTCG 378
 DB 244 IleThrTyrAspArgAla-----ValLysIleSerAspIleGlyThrSer 258
 QY 379 AGGCAG---TCATTCATGAGAGCGCGCTAGCGCTGAGGCGACCTCTGCGTCAAGGCC 435
 DB 259 LysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThrValAlaTrpMetAla 278
 QY 436 CCAAGATTCAGGCGCTGCATGTTATGATGAGAAGATGATGTTCTCTTGGAAATG 495
 DB 279 ProGluValIleArgAsnGluProValSerGluLysValAspIleThrSerPheGlyVal 298
 QY 496 GTGCTTACAGATGTTGCTGTGACAGACAGCGCCT----- 528
 DB 299 ValLeuThrProIleuLeuThrGlyGluIleProTyrLysAspValAspSerSerAlaIle 318
 QY 529 -----GCAGTGGCGACACACACAGCTCCAGATGCCAAGAGCTGTCCAAGGGCATCCG 582
 DB 319 IleTrpGlyValAlaGlySerAsnSerLeuHisLeu----- 329
 QY 583 CCGGTTCTCTGGGAGCGCGAGAGAGTGCAGTTCGGCGAGCTGCAGGCGCTCATGATGGAG 642
 DB 330 ProValProSerSerCysProAsp-----GlyPheLysIleLeuLeuAlaArg31n 345
 QY 643 TGGTGGACACTAGACAGAGAGAGCGACCG-----CTGCGCCCTGTG 664
 DB 346 CysThrAsnSerLysProAlaGlnAlaProSerPheArgGlnIleLeuLeuHisLysAlaP 365
 QY 645 GTGGTACCGACAGATGAAGACCGGACTTTGCCAGCTTCAAGTATGACAGTGTCTGGG 744
 DB 366 IleAlaAspAla-----AspValLeuSerThr 374
 QY 745 AAGCAGACAGCCTTCTTTCATCCAGGCGCAG-----GAGTAC 783
 DB 375 ProGlnGluThrTyrPheLysSerGlnAlaGluTrpArgGluGluValLysLeuHisPhe 394

QY 784 ACCGTGCTTTTGGATGCA----- 804
 DB 395 GluLysIleLysSerGlnGlyThrCysLeuHisArgLeuGluGluLeuValMetArg 414
 QY 805 ---AAGAGAGATCCAGAAC----- 822
 DB 415 ArgArgGluGluLeuArgHisAlaLeuAspIleArgGluHisTyrGluArgLysLeuGlu 434
 QY 823 -----TACAGGTGTGACACACA----- 843
 DB 435 ArgAlaAsnAsnLeuLysTyrMetGluLeuAsnAlaLeuMetLeuGlnLeuGluLysGlu 454
 QY 844 AAGGCGCTCAGGAGGCGCAGAGATGTGC-----TGCCGTGGAGTAAGGTGAGC 894
 DB 455 ArgGluLeuLeuArgArgGlnAlaLeuGlnAlaArgCysProGlyLeuLysLysP 474
 QY 895 TGCCAGCTCCAGGTCCAGAGATCCCTGTGACAGCCAGCAGACAGCAAAATCTACATC 954
 DB 474 OH1ProSerArgGlyLeuLeuHisGlyAsnThrMetGluLysLeuIleLysLysArg 494
 QY 955 TACACCTCAAGGCGATGTCCCTTAACACACACCCACAGGCTTGGATTACTCCAGCT 1014
 DB 494 nValProGlnAsnLeuSer-----ProHisSerGlnArgProAspIleLeuLysAl 511
 QY 1015 GTGCTACGCTCTTGTGGCGCTGCTGTTATTAAAGAATCTACCTGCTTAAGCG 1074
 DB 511 ArgLys-----LeuLeuProLysLeuAspAlaAl 521
 QY 1075 GCGCTCGCGAGATGGAGTGTGCTGTGTTCCGCTGCGGCGGACCCCAAGAGACAGC 1134
 DB 521 AleuSerGlyValGlyLeu----- 527
 QY 1135 TGCTCTACCTGCTGTCACACACAGCCACAGCTCCAGTTCAGCTCCGGATGAGAC 1194
 DB 528 ---ProGlyCysProLysAlaProProSer---ProGlyArgSerArgArgLysThr 545
 QY 1195 GCAGGCGAGACCCCTACCCAGTGAAGGCGCATGAGGTGTCAACGCGCTGTGAGGTC 1254
 DB 545 ArgHisArg-----LysAlaSerAlaLys 553
 QY 1255 TGGTACAGCAATGGCGCGCGCTCTGTATCAGTGCAGTGCCTCCGAGATGCACAG 1314
 DB 553 SerLysSerGlyAspLeuProGlyLeuArgThrAlaValProThrHisGluProGlyAl 573
 QY 1315 CCGCTGAGCGCT-----ACATGGCCC-----CTTCAGT 1344
 DB 573 yProGlySerProGlyGlyLeuGlyGlyGlyProSerAlaTrpGluAlaCysArgProAl 593
 QY 1345 GTTACGT-----CA 1353
 DB 593 AleuAlaGlyLeuHisHisAspLeuLeuArgLysMetSerSerSerProAspLe 613
 QY 1354 GTGCTGTGAGCTCTGAGGCGAGAGAGGAGTGTGCTGGGCGCTCG-----ATGAC 1407
 DB 613 uLeuSerAlaAlaLeuGlySerArgGlyArgGlyAlaThrGlyGlyAlaGlyAspProG 633
 QY 1408 AAGGCCAATCCTTGGTGTATGTAACACTCCACACCTTACAGCTGTGTCCCGGACTTTC 1467
 DB 633 ySerProProAlaArgLysAspThrProProSerGlnGlySerAlaProGlySerTh 653
 QY 1468 TGGG-----GGTTCCTCAGGCGCTTACGAGACATGTTCC 1503
 DB 653 rSerProAspThrProGlyGlyAlaLysGlyLysLysProPro-----Fr 668
 QY 1504 GTAGCGCTCTTGAGACAGGAGAC----- 1549
 DB 668 oValGlyPro-GlyGluGlyValAlaGlyLeuLeuGlyThrGlyArgGluGlyThrSerGlyA 688
 QY 1527 -----CCCGGACAGCCACACGCGCC----- 1548
 DB 688 rGlyGlySerArgAlaGlySerGln-HisLeuThrProAlaAlaLeuLeuTyrArgAla 707

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 15, 2003, 07:07:27 ; Search time 18.5 Seconds
(without alignments)
6946.990 Million cell updates/sec

Title:	US-09-836-392-8_COPY_22.2205
Perfect score:	4058
Sequence:	1 atcgtgagcaccctggc.....cttgacctgcagaagaag 2184

Scoring table:	
BLOSUM62	
Xgapop 10.0	Xgapext 7.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 525148

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frames+2dp_model_DEV=xlh
-o=/cgn2.1/USPO/spool/USO9896392/runat.08042003.090329.22829/app.query.fasta_1.2375/
-DB-issued Patents_AA_QFMT=fasta_n_SUFFIX=n2p_ral_MINMATCH=0.1-LOOCL=0
-LOOEXT=0-UNITS=bits-START=1-END=1-MATRIX=blomsum-TRANS=humana0.cdi
-LIST=45-DCCALIGN=200-THR_SCORE=pt-THR_MAX=100-THR_MIN=0-ALIGN=15
-MODE=LOCAL-OUFFM=ptc-NORM=ext-HEAPSIZE=500-MAXLEN=2000000000
-USER=USO9896392.ecgn.1.12.urnat.08042003.090329.22829-NCPU=6-ICPU=3
-NO_XLPPX-NO_MMAP-LANG=EUCYR-NEG_SCORES=0-WAIT-LONGIO-DEV_TIMEOUT=120
-WARN_TIMEOUT=30-THREADS=1-XGAPOP=10-XGAPEXT=0.5-FGAPOP=6-FGAPEXT=7
-YGAPOP=10-YGAPEXT=0.5-DELOP=6-DELEXT=7
```

Database :

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep: *
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep: *
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep: *
- 6: /cgn2_6/ptodata/1/1aa/Backfilltest.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	295.5	7.3	859	1	US-08-395-560-2	Sequence 2, Appl1
2	295.5	7.3	859	5	PCT-US95-02792-2	Sequence 2, Appl1
3	267	6.6	668	1	US-08-205-018-2	Sequence 2, Appl1
4	261	6.4	1050	4	US-09-428-711A-16	Sequence 16, Appl-1
5	245	6.0	1051	4	US-09-428-711A-14	Sequence 14, Appl-1
6	244	6.0	800	2	US-08-469-537A-72	Sequence 72, Appl-1
7	244	6.0	800	2	US-08-469-537A-78	Sequence 78, Appl-1
8	244	6.0	937	2	US-08-469-537A-105	Sequence 105, Appl-1
9	243.5	6.0	505	1	US-08-223-616-20	Sequence 20, Appl-1
10	243.5	6.0	505	4	US-08-446-668-20	Sequence 20, Appl-1
11	243	6.0	505	5	PCT-US95-04228-20	Sequence 20, Appl-1
12	243	6.0	455	3	US-09-221-235-5	Sequence 5, Appl1

13	243	6.0	455	3	US-09-221-988-5	Sequence 5, Appl 1
14	243	6.0	455	4	US-09-221-537-5	Sequence 5, Appl 1
15	243	6.0	455	4	US-09-221-236-5	Sequence 5, Appl 1
16	243	6.0	455	4	US-09-221-416-5	Sequence 5, Appl 1
17	243	6.0	455	4	US-09-221-245-5	Sequence 5, Appl 1
18	243	6.0	455	4	US-09-163-115-5	Sequence 5, Appl 1
19	243	6.0	455	4	US-09-121-528-5	Sequence 5, Appl 1
20	243	6.0	455	4	US-09-593-553-5	Sequence 5, Appl 1
21	243	6.0	455	4	US-09-221-237-5	Sequence 5, Appl 1
22	241.5	6.0	506	4	US-08-426-509A-6	Sequence 6, Appl 1
23	241.5	6.0	511	5	PCT-US-09-05008-6	Sequence 6, Appl 1
24	238.5	5.9	1037	4	US-09-428-711A-1	Sequence 21, Appl 1
25	236	5.8	263	3	US-09-035-706-5	Sequence 5, Appl 1
26	236	5.8	263	3	US-08-955-841-5	Sequence 5, Appl 1
27	236	5.8	263	4	US-09-390-422-5	Sequence 5, Appl 1
28	236	5.8	263	4	US-09-566-906-5	Sequence 5, Appl 1
29	236	5.8	831	1	US-07-928-464-2	Sequence 2, Appl 1
30	236	5.8	831	1	US-08-003-311B-2	Sequence 2, Appl 1
31	236	5.8	821	1	US-08-261-433-2	Sequence 2, Appl 1
32	236	5.8	821	5	PCT-US93-07347-2	Sequence 2, Appl 1
33	233	5.7	795	3	US-09-031-563-23	Sequence 23, Appl 1
34	233	5.7	795	4	US-09-392-277-23	Sequence 23, Appl 1
35	232	5.7	579	4	US-09-529-279-4	Sequence 4, Appl 1
36	232	5.7	560	4	US-09-529-279-15	Sequence 15, Appl 1
37	228	5.6	266	2	US-08-701-191A-9	Sequence 29, Appl 1
38	225	5.5	1052	3	US-08-863-118-2	Sequence 2, Appl 1
39	225	5.5	1317	3	US-09-083-521-7	Sequence 107, Appl 1
40	224.5	5.5	943	2	US-08-469-537A-107	Sequence 107, Appl 1
41	224	5.5	848	4	US-08-976-025-10	Sequence 10, Appl 1
42	223.5	5.5	1564	4	US-09-457-040B-27	Sequence 27, Appl 1
43	221	5.4	973	1	US-08-162-809-8	Sequence 8, Appl 1
44	220.5	5.4	612	2	US-08-673-789-11	Sequence 11, Appl 1
45	220	5.4	1276	1	US-08-222-610-24	Sequence 24, Appl 1

ALIGNMENTS

RESULT 1
 US-08-395-580-2
 Sequence 2, Application US/08395580
 Patent No. 5676945
 GENERAL INFORMATION:
 APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
 APPLICANT: Hospital of Philadelphia
 TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid
 TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereto
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock Washburn Kurtz MacKiewicz and No. 5676945rls
 STREET: One Liberty Place - 46th Floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: U.S.A.
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch disk, 720 Kb
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Wordperfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/395,580
 FILING DATE: herewith
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/205,018
 FILING DATE: 01-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Rebecca L. Ralph (formerly Gaumond)
 REGISTRATION NUMBER: 35,152
 REFERENCE/DOCKET NUMBER: CH-0488
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-568-3100
 TELEFAX: 215-568-3439

QY	963	CAGG-----	GCATGAGCCCTTAACACACCCCAAGGCTT	1001
Db	470	erAspPheGlnHisLysPheThrValGlnAlaSerProHMetAspLysArgLysSerL		490
QY	1002	GCATACCTCAGCTGTCGTCACCTGCTTCTTGCCGCGCTGTATA-----	TTAAAGAGA	1055
Db	490	eulIaAsnSerArgSerSerProProAlaSerProThrIleLeuProAlaGluAlaAl		510
QY	1056	TT-----	CCTACCTGCTCTTACGCGGCTCGCCGATG----	1087
Db	510	legIleuThrProGlyGlnSerSerLysThrTrpLys-----	ArgSerValValProL	529
QY	1088	-----	GCCTTCTGC-----	1096
Db	529	ysGIuGIuGIuGIuGIuGIuLysArgAlaProLysLysLysGlyArgThrTrpGlyP		549
QY	1097	-----	CTGTGTCCTCCGCTGTCGCGGCGACCCCAAGGACAG	1133
Db	549	roGIuYThrLeuGIuLysGluLeuAlaSerGIuLysProGlySerProGlnArgArg		569
QY	1134	CTGCTCTTACCTGTGCTCAGACACAGCCAAAGGTCCTCAAGTCAGATCGCGAGTGA		1193
Db	569	IuLysAlaAsnGlyLeuSerThrProSerGIuSerProHisPhe-HisLeuGIuLys		588
QY	1193	-----		1193
Db	589	serLeuValAspGIuYTrpLysGlnTrpSerSerSerAlaProAsnLeuValLysGlyPro		608
QY	1194	CGCAGCGCAGAACCCCTTACC-----		1214
Db	609	ArgSerSerProAlaLeuProGIuPheThrLysLeuMetGIuMetAlaLeuAlaAla		628
QY	1215	-----	AGTAAAGCCATGGAGTGT	1235
Db	629	serTrpValValProLeaSplLeuGIuLysProLysAspSerGIuLysProGlySerGIy		648
QY	1236	CACACAGCGCTCTGAGGTGGTACAGCAATGGCGCGGCTCTTTCATCCACTGCG		1255
Db	649	GlnHisSerProSerGlnSerTrpLysLys		662
QY	1296	CTCCCT-----	GGAGATCTGACGCGGCTGGAGCCCTACAT-----	1331
Db	663	IleProPheProArgGIuLysAspGIuLysProSerSerAspGIuLysIleHisGIuL		682
QY	1332	---GGCCCCCTCCAGTTCAGTCAGTCGTCAGCTCGAGG-----	CAG	1376
Db	683	ProThrProValAsnSerAlaThrSerThrProGlnLeuThrProHAsnSerLeuLys		702
QY	1377	AGGGAGAGAGGTGCTGCTGCTCGATGACAAAGCCAACTCCTTGGTGATGTACCACTC		1436
Db	703	ArgGIuGIuAla-----		706
QY	1437	CACCACCTTACCACCTGTGTCGCGGTACTTCTGCGGGGT-----		1475
Db	707	HisHis-----	ArgArgCysGIuValAlaLeuLeuGIuLysGIuAlaValLeuAlaAla	724
QY	1475	-----		1475
Db	725	ThrGIuLeuGIuLysPheAspLeuLeuGIuAlaGIuLysCysGlnLeuLeuProLeuGIuL		744
QY	1476	CCCCAGCCCCCTCAGGACATGTTCCCGTCGCGCCCTTGACA-----	CGG	1522
Db	745	ProGlnProProAlaArgGIuLysLysArgArgGIuLeuPheGlnArgSerSer		764
QY	1523	AACCCCGGAGCGACGACAGCGCCAAACCAAGG-----		1588
Db	765	ArgProArgArgSerThrSerProProSerArgLysLeuPheLysLysGIuLysProMet		784
QY	1559	TGCTGAGGGAGACTCCATCGCGGACGTGAGCATGTACAGTGAAGAGCTGGGACGC		1618
Db	785	LeuLeuLeuLysProSer-----	AlaSerLeuThrLeuLeuSerLeuSerSer	801
QY	1619	AGATCTGATCACACAGAAATCACTCACTG-----		1648

Db 802 ILSErGLuCYsAnSerThrArgSerLeuLeuArgSerAspSerAspLluLeValVal 821

QY 1649 -----ACTAGTCTTCATGTCCTCTACTCT 1675

Db 822 TyGluMetProValSerProValGluAlaProProLeuSerProCysThrHisAnPro 841

QY 1676 -----CATCCCAACCCCGGCGAGTCCAGCCGCCCTCAAGCCGCCACAGTCCCGG 1728

Db 842 LeuValAsnValArgValGluArgPheLysArgAspProAsnGlnSerLeuThrProThr 861

QY 1730 CAAGTTCCTCCAGTGCCTTTCTCCACGCAGTCCAGAGCAGCAGTCAATACATACCC 1788

Db 862 HisValThrLeuThrThrProSerGlnProSerSerHisArgArgThrProSerAspGly 881

QY 1790 CCGGTGTCCTCCGCGAGAGTCTGAGCATGACCTGACCC-----CCA 1833

Db 882 AlaLeuLysProGluThrLeuAlaSerArgSerProSerSerAsnGlyLeuSerPro 901

QY 1832 TGGAGGGGAGACCTTCAGCCAGCAGCAGCAGCAGCCGGAATGATCCAGCGGTGAGAGACC 189

Db 902 SerProGluValArgLysMetLeuLysThrProSerProSerArgAspProGlyLuhPhePro 921

QY 1892 TCATTTGGGTCCCGAGCGCGGTGGAGATGTATACGTACATGAGCCCTGAGAGA----- 1944

Db 922 ArgLeuProAspProAsnValValPheProProThrProArgArgThrThrGlnGln 941

QY 1943 -----AGGATTCGTGAAGCCCGAGCGGGGCGGAGTCATTCGCTTAAAGCCCGAGACC 199

Db 942 AspSerThrLeuGluArgProGlyThrLeuGluLuhPheLeuProArgProArgProSerAla 961

QY 1997 TGATCCCGCATNGGGGTGCTGTGATCTCTCCGTCGTGGCGCAAGACACGTGTGTGCA 205

Db 962 AsnArgGluArg-----LeuAspProThrTrp-----PheValSer 973

QY 2057 CCTTTGAATGAAGAAACAGACAGTGTGGTGGCCGCTGTGAGGGGCGTGGCCAGGG 2111

Db 974 ProSerHisAlaArgSerThrSer-----ProAla 983

QY 2117 AGTTGACATTTTTCACCAAGTCCGACG 2143

Db 984 AnSerSerSerThrGluThrProSer 992

RESULT 12

Q9W313 PRELIMINARY; PRT; 1020 AA;

AC Q9W313; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE CG2272 Protein.

GN MK2 OR CG2272.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;

CC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RX MEDLINE=20196006; PubMed=10731133;

RA Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Workman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Adrill J.F., Agdayani A., An H.-J., Andrews Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Boriva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,

RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

```

Db 306 AlatyAsnGlnIynValSpValTyrSerPheGlyLeuIprGluLeuIleThr 325
QY 517 GGACAGCGCCCTGCTACTGGGCCACCACTCCAGATTGCC---AAGAAGCTGCCAAG 573
Db 326 GlyLeuLeuProPheGlnAsnMetThrAlaValGlnAlaLeuAlaValAsnArg 345
QY 574 GGCATCCGCCCGTCTGGGGCAGCCGAGAGAGTCCAGTCCGGCGACTGCAGGCGCTC 633
Db 346 GlyValArgProThrVal-----ProAsnAspCys---LeuProValLeuSerAspIle 362
QY 634 AGCATGAGGCGCGGAGCACTAGCAGAGAACGACCGCTGCCCTGCTCGGTGGAGC 693
Db 363 MetThrArgCysTrpAspAlaAsnProGluValArgProCysPheValGluValIys 382
QY 694 CAGATGAG 702
Db 383 LeuLeuGlu 385

RESULT 11
Q9H2N5 PRELIMINARY; PRT; 1066 AA.
ID 09H2N5;
AC 09H2N5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Mixed lineage kinase MLK1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA McNeae J.J., Dower S.K., Guesdon F.;
RT cDNA sequence and gene organisation of mixed lineage kinase 1.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC 1. SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL: AF251442; AAG44591.1; -.
DR HSSP: P29355; ISEM.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00326; SH3; 1.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50002; SH3; 1.
DR KINASE: SH3 domain.
FT NON_TER
SQ SEQUENCE 1066 AA: 118463 MW: EDD08EBEEF7482723 CRC64;

Alignment Scores:
Pred. No.: 7, 82e-14 Length: 1066
Score: 301.50 Matches: 208
Percent Similarity: 32.81% Conservative: 104
Best Local Similarity: 21.87% Mismatches: 323
Query Match: 7.43% Indels: 318
DB: Gaps: 38

US-09-836-392-8_COREY_22_2205 (1-2184) x Q9H2N5 (1-1066)
QY 7 AGGCACCTGCGGCGCCAGCATGCCATGAGAACTCTCCGAGTTCGGCGAGAGGCCAGC 66
Db 122 ArgHisAspProAspGluAspIleSerGlnThrIleGluAsnValArgGlnGlnAlaIys 141

```

```

QY 67 ATGCTCAGCGCGCTGAGCAGACCCCTGCATCTGCGCCTCATCGCATCAGATCCACCG 126
Db 142 LeuPheAlaMetLeuIynShsProAsnIleIleAlaLeuArgIleValCysLeuIysGlu 161
QY 127 -----CTGCTGCTGCCCTGAGAGCTCGCGCGCTCAGACAGCTCAACACCGCTGCTCC 180
Db 162 ProAsnLeuCysLeuValMetGluPheAlaArgIleGlyLeuProLeuAsnArgValLeuSer 181
QY 181 GAGACGCCAGAGATCTCTCTTATACCCCTGGAGACATGCTCAACCCAAAATA--- 237
Db 182 -----GlyLysArgIleProProAspIleLeu 190
QY 238 -----GCCATCAGATCGCGCTGCGCGCTGCCCTGACGAGAA----- 279
Db 191 ValAsnTrpAlaValGlnIleAlaIleArgGlyMetAsnTyrLeuHisAspGluAlaIleVal 210
QY 280 AACATCATCTTCTGTGACCTGGAAGTGGACACATCTTGCTGTGTCCTT----- 330
Db 211 ProIleIleHisArgAspLeuIysSerAsnIleLeuIleLeuIynValGluAsn 230
QY 331 ---GAGCTCAGAGGACACATCAACATCAAGCTATCGACTACGAGGATTCGAGGAGTCA 387
Db 231 GlyAspLeuSerAsnIysIle---LeuLysIleThrAspPheGlyLeuAlaArgGluTrp 249
QY 388 TTCATGAGAGGCGCCCTAGGCGGTGAGGCGACCTGCTGCTACGAGCCAGATCAGG 447
Db 250 HisArgThrThrIysMetSerAlaAlaGlyThrTyrAlaTrpMetAlaProGluValIle 269
QY 448 CCTGCATTGTATATGATGAGAGAGATGATATGCTCTATGAAATGCTCTACAG 507
Db 270 ArgAlaSerMetPheSerIysGlySerAspValTrpSerTyrGlyValLeuLeuTrpGlu 289
QY 508 TTGCTCTCAGGAGAGCGCCCTGACTGGGCCACACAGCTCCAGATTGCC----- 558
Db 290 LeuLeuThrGlyGluValProPheArgGlyIleAspGlyLeuAlaIleArgIleVal 309
QY 559 -----AAGACCTGTCCAAAGGCGATCCCGCTGCTGGGCGAGCGGAGAGTGCAG 612
Db 310 AlaMetAsnIysLeuAlaLeuProIleProSerThrCysProIleuro----- 325
QY 613 TTCGGCGCACTGCGAGCGCTCATGATGAGAGTGTGGACACTAAGCCAGAGACAGCAG 672
Db 326 -----PheAlaLysLeuMetGluAspCysTrpAsnProAspProHisSerArg--- 341
QY 673 CTTGGCCCTGTGCTGTGAGCCAGATGAGAGACGACTTTGCCACCTTCATGATGA 732
Db 342 -----ProSerPheThrAsnIleLeuAspGln 350
QY 733 CTGTGCTGTGGAGAGACAGACGCTTCTTC-ATCCAGAGGCCAGAGATACACGCTGT 791
Db 351 LeuThrThrIleGluIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 370
QY 792 GTTTTGGATGGAAGAGAGAGAGTCCAGACTACAGGT-----GATGACACAGA 842
Db 371 GlnAspAsnTrpIysHisGluIleGlnIleuMetPheAspGlnLeuArgAlaIysGluLys 390
QY 843 GAA-----GGGCTCATGAGAGTGCAGAG 866
Db 391 GluLeuArgThrTrpGluGluIleuThrArgAlaAlaLeuGlnIleuAsnGlnIleu 410
QY 867 GATGTGCTCCCTCGGAGAGAGAGTGCAGGCTCCAGAGTCCAGAGAT----- 916
Db 411 GluLeuLeuArgArgArgIleGluIleuLeu---AlaGluTrpGluIleAspIleLeuGlu 430
QY 917 -----CCCTGTGAGAGAGCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
Db 430 ArgGluLeuAsnIleIleIleHisGlnLeuCysGlnIleuLysProArgValIysLysArg 450
QY 947 -----TTACATCTACACCTT 962
Db 450 yscGlyLysPheArgLysSerArgLeuLysAspLysAspLysAsnArgIleSerLeuPro 470

```



```

Db 368 -----ProApPhealaserlleuengln 376
Qy 733 CTGTG-CTGTGGAGACAGACACCTTCTTCATCCAGGGCCAGGATACCGGTGT 791
Db 377 LeuGluAlaLeuGluAlaValLeuArgLumelProArgAspSerPheHisSerMet 396
Qy 792 GATTGGATGGAAAAGAGAGAGTCAG-----GAATACACGGTGGTGAACAC 839
Db 397 GlnGluGlyTyrPyrAspGluIleGlnGlyLeuPheAspGluLeuArgAlaLysGlu 415
Qy 840 AGAAGAGGGCTCATAGAGATGAGATGTGCTGCTGGATGAGATGAGATGAGATGAG 899
Db 416 -----LysGluLeuLeuSerArgGluGlu-----GluLeuThr 426
Qy 900 GCTCCAGCTCCAGAGATCCCTGTGGACACCCAGCCAGGACAGAAATCTACTCTAC 959
Db 427 ArgAlaAlaArgGluGluArgSerGlnAlaGluGlnLeuArgArgGluGlnLeu 446
Qy 960 CCTCAAGGCGATGTGCGCCCTTAACACACCCCAAGCCCTTGATACACAGCTGT 1016
Db 447 AlaGlnTyrGluLeuGluAlaPheGluArgAspValThrLeuLeuGlnGlnValAsp 466
Qy 1017 -----CGTCACTGCTTCTTGGCGCTGCTTATTAAGAAATCTCACTGTGT 1067
Db 467 ArgGluArgProHisValArgArgArgArgGlyThrPheLysArgSerLysLeuArgAla 486
Qy 1068 CTTAGGGGGCTCCCGGATGGCTGTGTGCTGTGCTTCCCGTGGTGGGCGACCCCAA 1127
Db 487 ArgAspGly----- 489
Qy 1128 GGACACCTGCTCTCTACCTGTGCTCAACACAGCAGCAGGATCCAGTTCAG---CATCG 1184
Db 490 GlyGluArgIleSerMetProLeuAspPheLysHisArgIleThrValGlnAlaSerPro 509
Qy 1185 GGATGAAGACGACGCGCAGAACCCCTACCCAGTGAAGGCC----- 1224
Db 510 Gly-LeuAspArgArgArgAspValPheGluValGlyAlaGlyAspSerProThrPheP 529
Qy 1225 -----ATGAGAGTGGTCAACACCGGCTGTGAGCTGTGAGTCAACGAA 1265
Db 529 oArgPheArgAlaIleGlnLeuInProThrGluSerGly---GlnThrTyrIleArg 548
Qy 1266 TGGGCGGCGCTCTCTG---TCATCAGCTGTGCTCTCTGAGATGTGACGGCGGTGA 1322
Db 548 nSerPro-ArgArgLeuGluAspSerSerAsnGlyLumArgAlaCysTyrAlaTyr 568
Qy 1323 GCCCTACATGGCCCTCCATGAGTGTACGTAGTGTGCAAGCTGTGAGGAGGAGGGA 1382
Db 568 LProSerSerProLysProGlyGluAlaGlnAsn---GlyArgArgArgSerArgMetAs 587
Qy 1383 GAGAGCTGTGCTGTGCTGAGTACAGCAAGCCAACTCTGTGATATACACCTCCACAC 1442
Db 587 pgLualAthrTyrPtyrLeuAspSerAspAspSerProLeuGlySerProSerThrP 607
Qy 1443 CTACACAGCTGTGCGCGGACTCTGTGCGGAGCCAGCCAGCCCTCAGGAGCATGTTCC 1502
Db 607 ro-----ProAlaLeuAsnGlyAsnProProArgProSer----- 618
Qy 1503 CGTGGCGCCCTTGGAGACAGCAACCCCGGAGCCAGCCAGCCAGCCAGCCAGG--- 1558
Db 619 -----ProGluProGluGluProArgArgAlaGly-----ProGlnGluArgGly 634
Qy 1559 -----TGCTGAGGGGAGCTCCA----- 1576
Db 634 snSerSerGlyThrProLysLeuIleGlnArgAlaLeuLeuArgGlyThrAlaLeuLeu 654
Qy 1577 -----TCGCGAGCTGAGCAT 1592
Db 654 laserLeuGlyLeuGlyArgAspLeuGlnProProGlyGlyLeuSerArg----- 670
Qy 1593 CATGTACAGTGAAGAGAGCGGAGCAGCAGATCTGTATCCAGCAGGAATCACTCACTGA 1652

```

```

Db 671 -----GluArgGlyGluSerProThrAs 678
Qy 1653 CTGTCCATGTCTCTCTACTCTCATFCCCAACCCCGGAGCTGCCAGTCCCT----- 1708
Db 678 LArProProProLagInMetProSerProCysProProGluLeuProSerThrProLeuI 698
Qy 1709 -----CAGCCGCCAGCTCCCGCCAGCAAGTCTT- 1738
Db 698 LArgLeuSerInThrThrProAspAlaHisSerSerProThrProGlyProLeuLeuL 718
Qy 1739 -----CAGTGTGCTTCTTCACAGCAGCTGAGAGACTC 1772
Db 718 euAspLeuGlyValProSerGlyGlnProSerAlaLysSerProArgArgGluGluThr 738
Qy 1773 AGACATGCTACATAGCCCGCGGTGCTCCCGACAGCTGTGAGCAAGCTGACCCCAT 1832
Db 738 rgGly-----ArgInValSerProProProGlyLleSerArgSerAlaProGlyT 755
Qy 1833 GAGAGGGAGACCTTCAGCCAGCAGCAGCTGCAGC-----CGTGAAGATCTCGCCGT 1883
Db 755 hrProGlyThrPro-ArgSerProProLeuGlyLeuIleSerArgProArgProSerPro 774
Qy 1884 CAGAGACTCATTTGGGTCCCGCAGCGCGGTGAGATGTTATCGTCAATTCGCTGAGAA 1943
Db 775 LeuArgSerArgIleAspPro-----Tyr----- 782
Qy 1944 GATTCCTGAAGCCCGGCGGCGGAGTCAATGCGCTTAAAGACCCGAGA----- 1994
Db 783 SerPheValSer---AlaGlyProArgProSerProLeuProSerProGlnProAlaPro 801
Qy 1995 -----GCTGACTCGCATGGGAGTGT-----GGT 2018
Db 802 ArgArgAlaProTyrPheThrLeuPheProAspSerAspProPheThrAspSerProProAla 821
Qy 2019 GGATGCTGCCGTGTGTCGCAAGACACTGT 2048
Db 822 AsnProPheArgGlyGlySerGlnAspCys 831

RESULT 9
Q9M085 PRELIMINARY; PRT; 412 AA.
AC Q9M085;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Protein kinase-like protein.
GN AT4G31170.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
RA Mews H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AL161576; CAB79835.1;
DR HSSP: P08631; IAD5.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; TYR_pkinase.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase.1.
DR SMART: SM00221; STYC; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

```


QY 1921 GTATCGCATGCGCTGAGAGAGATTCTGAAGCCAGCGGGCGAGTCATTGCCCTC 1980
 DB 1282 VALLIIEVALLIEGLIENGLINUSPSPERGIAAGIARGLYARVALIALEALVAL 1301
 QY 1981 TTTAAAGCCGAGAGACTGATCCGATGGGTGCTGTGATGCTCCCTGTGGCAAG 2040
 DB 1302 LEULYSALARGJULENTHRPHONISGLYVALLEVALSALALVALVALALALYS 1321
 QY 2041 GACACTGTGTGTGACACTTTGAAAATGAAAACACAGAGTGTGCTGCGCTGTGGAG 2100
 DB 1322 ASPHTRVALVALCYSTHRPHEGLINUSNGLIUSNTHRGITRIPCYSLEULIATRPARG 1341
 QY 2101 GCGTGGGGCGGAGAGAGTGCACATTTCTACAGTCTTACAGAGAGTGGCGGCTG 2160
 DB 1342 GLYTRPGIYALARGJULPHASPILEPHETRYGLINSETRYGLINLEUGLIYARGL 1361
 QY 2161 GAGCTTGCACTGCCAAGAGAG 2184
 DB 1362 GIUALACYS THRARGLYSARGARG 1369

RESULT 2

Q9H5S3 PRELIMINARY: PRT: 253 AA.
 ID Q9H5S3;
 AC 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CDNA: FLJ23119 f1s, Clone LNG07978.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
 RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
 RA Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y., Ota T., Suzuki Y.,
 RA Obaraishi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
 RA Iisogai T., Sugano S.;
 RT "NEO human cDNA sequencing project."
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK026772; BAB1547.1; -
 SQ SEQUENCE 253 AA: 27913 MW: FARD58E9E29FE741 CRC64;

Alignment Scores:

Pred. No.: 3,28e-90 Length: 253
 Score: 1325.00 Matches: 252
 Percent Similarity: 99.60% Conservative: 0
 Best Local Similarity: 32.65% Mismatches: 1
 Query Match: 32.65% Indels: 0
 DB: 4 Gaps: 0

US-09-836-392-8_COPY_22_2205 (1-2184) x Q9H5S3 (1-253)

QY 1426 ATGTACCACTCCACCACTACCAAGCTGTGCGCCGACTTCGGGGGTCCCAAGCCC 1485
 DB 1 MetTyHisSerThrThrThrGlnLeuGlyAlaArgThrPheCysGlyValProSerPro 20
 QY 1486 CTGAGGACATGTTCCCGTGGCGCCCTTGAGACAGGAAACCCCGGAGCCAGCCACAG 1545
 DB 21 LeuArgAspMetThrProValArgProLeuAspThrGlnProAlaAlaSerHisThr 40
 QY 1546 GCCAACCCAAAGTGGCTGAGGGGAGCTCCATCGCGAGCTGAGCATGATCACTGAG 1605
 DB 41 AlaAspProGlyValProGlnGlyAspSerIleAlaAspValSerIleMetTySerGln 60
 QY 1606 GACCTGGGCGAGAGATCCGATCCACAGAGATCACTACTACTCTGCTCCATGCC 1665
 DB 61 GlnLeuGlnThrGlnIleLeuIleHisGlnGlnSerLeuThrAspTrpCysSerMetSer 80
 QY 1666 TCTACTCTCATCCCAAGCCGAGGCTGCCAGTGTCCCTTCACAGCTCCAGCTCC 1725

DB 81 SerTySerSerSerProProArgGlnAlaAlaArgSerProSerSerLeuProSerSer 100
 QY 1726 CCAGCAAGTCTTCAGTGTGCTTTCCACCGAGCTGGAGACTGCATGCTACTAC 1785
 DB 101 ProAlaSerSerSerSerValProPheSerThrAspCysGlnAspSerAspMetLeuHis 120
 QY 1786 ACGCCCGGTGCTGCTCCGACAGAGTCTGAGCATGACCTGACCCCATGGAGCGGAGACC 1845
 DB 121 ThrProGlnAlaAlaSerAspArgSerGlnHisAspLeuThrProMetAspGlyLeuThr 140
 QY 1846 TTCAGCCAGCACTGAGCGCGGTGAAGATCTGCGCGTCAGAGACCTCATTTGGTCCC 1905
 DB 141 PheSerGlnHisLeuGlnAlaValLysIleLeuAlaValArgAspLeuLeuTrpValPro 160
 QY 1906 AGCCGCGGTGAGANGTTCATTCGCTGAGGAGAAAGATTCTGAACCCAGGCGGGCC 1965
 DB 161 ArgArgGlyGlyAspValIleValIleGlyLeuGlnLysAspSerLeuValGlnArgGly 180
 QY 1966 CGAGTCATGCGCTCTTAAAGCCGAGAGTGAAGTCCGATGCGGAGGTGCTGTGATGCT 2025
 DB 181 ArgValIleAlaValLeuLysAlaArgGlnLeuThrProHisGlyValLeuValAspAla 200
 QY 2026 GCGGTGTTGGCAAGGACACTGTGTGTCACCTTTGAAAATGAAAACACAGAGTGTGC 2085
 DB 201 AlaValValAlaLysAspThrValValCysThrPheGlnAsnGlnAsnThrGlnTrpCys 220
 QY 2086 CTGGCGCTGTGAGGGGCTGGGGCGCCAGAGAGTTGACATTTTTCACAGATCTACGAG 2145
 DB 221 LeuAlaValAlaTrpArgGlyTrpGlyAlaArgGlnPheAspIlePheTrpGlnSerTrpGln 240
 QY 2146 GAGCTGGCGGCTGAGGCTTGACCTGCCAAGAGAG 2184
 DB 241 GlnLeuGlnArgLeuGlnAlaCysThrArgLysArgArg 253

RESULT 3

Q8R085 PRELIMINARY: PRT: 252 AA.
 ID Q8R085;
 AC 08R085;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Similar to hypothetical protein FLJ23119.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RA Strausberg R.;
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC027199; ANH27199.1; -
 SQ SEQUENCE 252 AA: 28042 MW: 8E9BBBC4F600E95F CRC64;

Alignment Scores:

Pred. No.: 1,73e-69 Length: 252
 Score: 1046.50 Matches: 204
 Percent Similarity: 84.58% Conservative: 10
 Best Local Similarity: 80.63% Mismatches: 38
 Query Match: 25.79% Indels: 1
 DB: 11 Gaps: 1

US-09-836-392-8_COPY_22_2205 (1-2184) x Q8R085 (1-252)

QY 1426 ATGTACCACTCCACCACTACCAAGCTGTGCGCCGACTTCGGGGGTCCCAAGCCC 1485
 DB 1 MetTyHisSerLeuThrThrThrGlnLeuGlyAlaArgThrPheCysGlyValProAspPro 20
 QY 1486 CTGAGGACATGTTCCCGTGGCGCCCTTGAGACAGGAAACCCCGGAGCCAGCCACAG 1545
 DB 21 LeuArgAspThrPheSerValGlnProSerValLeuGlnThrPro--GlySerHisLys 39

DR InterPro: IPR001005; Myb_DNA_binding.
 DR Pfam: PF00069; Kinase; 1.
 DR ProDom: PD000001; Euk_kinase; 1.
 DR PROSITE: PS00037; MYB_1; UNKNOWN; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR Repeat: KD repeat.
 FT NON_TER 1
 SQ SEQUENCE 1369 AA: 153626 MW: AE90CE09C88370D9 CRC64;

Alignment Scores:

Pred. No.: 9.69e-276 Length: 1369
 Score: 3822.00 Matches: 728
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 94.18% Indels: 0
 Gaps: 0

US-09-836-392-8_COPY_22_2205 (1-2184) x Q96UN5 (1-1369)

QY 1 ATGCTGAGGACACCTGCGGCGCCACCGATGCGATGAAGAATTCTCCGAGTTCCGCGAGAG 60
 DB MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPhSerGluPheArgInglu 661
 QY 61 GCCACATGCTGCGCGCGCTGCGACACCCCTGCGATGCGCGCTATGCGCATCAGATC 120
 DB AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle 681
 QY 121 CACCGGCTGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 DB HisProLeuCysPheAlaLeuGlnLeuAlaProLeuSerSerLeuAsnThrValLeuSer 701
 QY 181 GAGAACGCCAGAGATTCCTTCTTATACCCCTGGGACACATGCTCACCAAAAAATAGCC 240
 DB GluAsnAlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnLysIleAla 721
 QY 241 TACCAAGATCGCTCGGCGCTGCGCTGCGACAGAAAGAAATATATCTTCTGTGACTG 300
 DB TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 741
 QY 722 TGTGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu 741
 QY 301 AAGTGGACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB LysSerAspAsnIleLeuValIleTrpSerLeuAspValLysGlnHisIleAsnIleLysLeu 761
 QY 742 LysSerAspAsnIleLeuValIleTrpSerLeuAspValLysGlnHisIleAsnIleLysLeu 761
 QY 361 TCTGACTAGGGGATTCGAGGAGTATTCATGAGGGCGCGCTGAGGGCGGACT 420
 DB SerAspTyrGlyIleSerArgIleSerPheHisGlnGlyAlaLeuGlyValIleGlyThr 781
 QY 421 CCTGCTACACAGGCG 480
 DB ProGlyTyrGlnAlaProGlnIleArgProAlaGlyIleValTyrAspGluLysValAspMet 801
 QY 782 ProGlyTyrGlnAlaProGlnIleArgProAlaGlyIleValTyrAspGluLysValAspMet 801
 QY 481 TCTGCTAGGAATGGTGTCTACGAGTGTCTGAGGAGCGCGCGCGCGCGCGCGCGCG 540
 DB PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnAlaTrpAlaLeuGlnHis 821
 QY 802 PheSerTyrGlyMetValLeuTyrGlnLeuLeuSerGlyGlnAlaTrpAlaLeuGlnHis 821
 QY 541 CACGAGCTCCAGATTGCCAAGAAGCTGTCCAAGGCGATCCGCCGCTTGGGGAGCGCG 600
 DB HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnPro 841
 QY 822 HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnPro 841
 QY 601 GAGGAAGTCAATTCGCGGCGACTGCGAGCGCGCTCATGATGAGTGTGGGACCTAAGCCA 660
 DB 842 GluGluValAlaGlnPheArgTyrLeuGlnAlaLeuMetMetGluCysTyrAspThrLysPro 861
 QY 661 GAGGAGCG 720
 DB GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 881
 QY 862 GluLysArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThr 881
 QY 721 TTCATGATGAAGTGTCTGTGGGAGACAGACAGCGCTTCTCATCCAGGCGCGAGAG 780
 DB PheMetTyrGlnLeuCysArgGlyLysGlnThrAlaPhePheSerSerGlnGlyInglu 901
 QY 882 PheMetTyrGlnLeuCysArgGlyLysGlnThrAlaPhePheSerSerGlnGlyInglu 901
 QY 781 TACACCGTGTGTTTGGATGGAAGAAAGAGAGTCCAGAACTACACGCTGTGGAACAA 840
 DB 781 TACACCGTGTGTTTGGATGGAAGAAAGAGAGTCCAGAACTACACGCTGTGGAACAA 840

DB 902 TyrThrValAlaPheTrpAspGlyLysGlnLysSerArgAsnTyrThrValAlaAsnThr 921
 QY 841 GAGAGGCGCTCATGAGAGTGCAGAGATGTGCTGCGCTGGATGAAGTGAAGCTGCCAG 900
 DB GluLysGlyLeuMetGlnValGlnArgMetCysCysProGlyMetLysValSerCysGln 941
 QY 901 CTCAGAGTCCAGAGATCCCTGTGACACCGCACCGAGACCGAAGAAATATACATCTACAC 960
 DB LeuGlnValAlaArgSerLeuTrpThrAlaThrGlnAspGlnLysIleTyrIleThr 961
 QY 942 LeuGlnValAlaArgSerLeuTrpThrAlaThrGlnAspGlnLysIleTyrIleThr 961
 QY 961 CTCAGAGGCGAGTGGCGCTTAAACACACCGCGCTGGAATCCGAGCTGTCTC 1020
 DB LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAla 981
 QY 982 LeuLysGlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValAla 981
 QY 1021 ACCTGCTCTTGGCGCGCTGCTGTTATTAAGAAATTCCTACTGCTGTGAGCGCGCTC 1080
 DB ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 1001
 QY 982 ThrCysPheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeu 1001
 QY 1081 GCCATGGGCTTGTGCGCTGTGCTCCGTGCGGTGCGGGCGACCCCAAGAGACGTGCTCC 1140
 DB AlaAspGlyLeuValAlaValAlaPheProValValArgGlyThrProLysAspSerCysSer 1021
 QY 1141 TACCTGTGCTCACACACAGCCCAAGCTCCAGAGTCCAGATCCGCGAGAGAGACGACG 1200
 DB TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 1041
 QY 1022 TyrLeuCysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspGluAspAlaArg 1041
 QY 1201 CAGAACCCCTACCCAGTGAAGGCCATGAGGAGGTGTCACAGCGGCTGTGAGTGTGTAC 1260
 DB GlnAsnProTyrProValLysAlaMetCylValAlaAsnSerGlySerGluValIleTrpTyr 1061
 QY 1042 GlnAsnProTyrProValLysAlaMetCylValAlaAsnSerGlySerGluValIleTrpTyr 1061
 QY 1261 AGCAATGGGCGCGCGCTCTTGTCTATCGACTGTGCTGCTCCCTGAGATGTGACGCGGCTG 1320
 DB SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgTyrLeu 1081
 QY 1062 SerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgTyrLeu 1081
 QY 1321 GAGCCCTACAGGCG 1380
 DB GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgGly 1101
 QY 1082 GluProTyrMetAlaProSerMetValThrSerValValCysSerSerGlnGlyArgGly 1101
 QY 1381 GAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 DB GluGluValAlaValTrpCysLeuAspAspLysAlaAsnSerLeuValMetCylHisSerThr 1121
 QY 1102 GluGluValAlaValTrpCysLeuAspAspLysAlaAsnSerLeuValMetCylHisSerThr 1121
 QY 1441 ACTTACAGCTGTGTGCGCGGCTACTTGTGCGGGGTCCCGACGCCCTGAGGACATGTTT 1500
 DB ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 1141
 QY 1122 ThrTyrGlnLeuCysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPhe 1141
 QY 1501 CCGCGGCGCGCTTGGACAGGAAACCGCGCGGAGCCAGCCACACAGCCCAACCAAGTGTG 1560
 DB ProValArgProLeuAspThrGlnIleProProAlaAlaSerHisThrAlaAsnProLysVal 1161
 QY 1142 ProValArgProLeuAspThrGlnIleProProAlaAlaSerHisThrAlaAsnProLysVal 1161
 QY 1561 CCTGAGGGGAGCTCCATGCGCGGAGCTGAGCATCATGATGAGGAGAGCTGGGCGAGCG 1620
 DB ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGlnLeuGlyThrGln 1181
 QY 1162 ProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGlnGlnLeuGlyThrGln 1181
 QY 1621 ATCTGATCCACAGGAATCACTCATCTACTGCTCCATGCTCTCTCATCTC 1680
 DB IleLeuIleHisGlnGlnLysSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 1201
 QY 1182 IleLeuIleHisGlnGlnLysSerLeuThrAspTyrCysSerMetSerTyrSerSerSer 1201
 QY 1681 CCACCGCGCGAGCTGCCAGAGTCCCGCTCAAGCTCCCGACAGTCCCGCGAGAGTCTTCC 1740
 DB ProProAlaArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 1221
 QY 1202 ProProAlaArgGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSer 1221
 QY 1741 AGTGTGCTTCTCCACGAGCTGCGAGAGTCCAGAGTCCAGATGCTCATACGCGCGCTCTCC 1800
 DB SerValProPheSerThrAspCysGlnAspSerAspMetLeuHisThrProGlyAlaAla 1241
 QY 1222 SerValProPheSerThrAspCysGlnAspSerAspMetLeuHisThrProGlyAlaAla 1241
 QY 1801 TCCGACAGTGTGACAGTACATGACCTGACCCCATGACGGGGAGACCTTGACGACGACTG 1860
 DB 1242 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 1261
 QY 1242 SerAspArgSerGlnHisAspLeuThrProMetAspGlyGluThrPheSerGlnHisLeu 1261
 QY 1861 CAGGCGGTGAAGATCCCTGCGCGTCCAGAGACCTCATTTGGTCCCGACGCGCGTGAAGAT 1920
 DB GlnAlaValLysIleLeuAlaValaLarGAspLeuIleTyrValProAlaArgGlyGlyLysP 1281
 QY 1262 GlnAlaValLysIleLeuAlaValaLarGAspLeuIleTyrValProAlaArgGlyGlyLysP 1281

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 15, 2003, 06:59:24 ; Search time 62 Seconds

(without alignments)
14516.352 Million cell updates/sec

Title: US-09-836-392-8_COPY_22_2205

Perfect score: 4058
Sequence: 1 atgtgagcagcactgcgggc.....cttgactcgcaagaagaag 2184

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODL=frame+ n2p model -DEV=rlh
-DB=SPTRMBL_21 -OFMT=fastan -SUFFIX=n2p -ISPT=MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum2 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR=SCORE=500 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09836392.8CGN.1.1.81.ernal.08042003.090328.22781 -NCPU=3
-NO_XLPPY -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriopl:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match Length	DB ID	Description
1	3822	94.2	1369	4	Q96JN5

2	1325	32.7	253	4	Q9H5S3	Q9H5S3 homo sapien
3	1046.5	25.8	252	11	Q8R0B5	Q8R0B5 mus musculu
4	1031	25.4	221	4	Q9BS11	Q9BS11 homo sapien
5	403.5	9.9	1175	5	Q9T2M4	Q9T2M4 caenorhabdi
6	364.5	9.0	2308	5	Q9VDJ9	Q9VDJ9 drosophila
7	337.5	8.3	847	4	Q16584	Q16584 homo sapien
8	334.5	8.2	850	11	Q9J015	Q9J015 mus musculu
9	315	7.8	412	10	Q9W085	Q9W085 arabidopsi
10	314	7.7	411	10	Q9Z031	Q9Z031 homo sapien
11	301.5	7.4	1066	4	Q9Z2M5	Q9Z2M5 drosophila
12	299.5	7.4	1020	5	Q9W313	Q9W313 chlamydomo
13	298.5	7.4	859	4	Q8W255	Q8W255 homo sapien
14	298.5	7.4	859	4	Q8W255	Q8W255 oryza sativ
15	296	7.2	1146	5	Q95VPE	Q95VPE drosophila
16	293.5	7.2	1161	5	Q95UN8	Q95UN8 mus musculu
17	293.5	7.2	1001	11	Q8VNG6	Q8VNG6 mus musculu
18	292	7.2	1036	4	Q07912	Q07912 homo sapien
19	291	7.2	1036	4	Q07912	Q07912 homo sapien
20	287	7.1	416	10	Q94C42	Q94C42 triticum ae
21	282	6.9	1036	4	Q8W2M1	Q8W2M1 homo sapien
22	278.5	6.9	546	10	Q22558	Q22558 arabidopsi
23	277.5	6.8	1637	6	Q9XSV8	Q9XSV8 bos taurus
24	277.5	6.8	5146	6	Q8SPM4	Q8SPM4 bos taurus
25	276.5	6.8	1055	11	Q54967	Q54967 mus musculu
26	273.5	6.7	981	15	Q92809	Q92809 abelson mur
27	271	6.7	411	10	Q9ANA6	Q9ANA6 arachis hyp
28	270.5	6.7	553	10	Q81808	Q81808 arabidopsi
29	270.5	6.7	1171	10	Q9SG74	Q9SG74 arabidopsi
30	269.5	6.6	570	10	Q8RML6	Q8RML6 arabidopsi
31	268.5	6.6	1040	5	Q8T720	Q8T720 caenorhabdi
32	268.5	6.6	438	10	Q9X187	Q9X187 arabidopsi
33	268.5	6.6	438	10	Q8S9J9	Q8S9J9 arabidopsi
34	268.5	6.6	445	10	Q9LME8	Q9LME8 arabidopsi
35	268	6.6	475	10	Q9SG75	Q9SG75 arabidopsi
36	267.5	6.6	1147	10	Q9SA26	Q9SA26 arabidopsi
37	266	6.6	978	5	Q8T721	Q8T721 caenorhabdi
38	264	6.5	1553	5	Q8T219	Q8T219 dictyosteli
39	263	6.5	422	5	Q23846	Q23846 dictyosteli
40	263	6.5	483	10	Q8RY96	Q8RY96 arabidopsi
41	263	6.5	527	5	Q9B125	Q9B125 dictyosteli
42	263	6.5	564	4	Q9H117	Q9H117 homo sapien
43	263	6.5	1029	10	Q23015	Q23015 arabidopsi
44	263	6.5	1338	5	Q23927	Q23927 dictyosteli
45	262	6.5	378	10	Q9L016	Q9L016 arabidopsi

ALIGNMENTS

RESULT 1

ID	Q96JN5	PRELIMINARY:	PRT: 1369 AA.
AC	Q96JN5		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	KIAI1790 protein (Fragment).		
GN	KIAI1790.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RX	MEDLINE=21245130; PubMed=11347906;		
RA	Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.,		
RT	"Prediction of the coding sequences of unidentified human genes. XX.		
RT	The complete sequences of 100 new cDNA clones from brain which code		
RT	for large proteins in vitro."		
RL	DNA Res. 8:85-95(2001).		
CC	-1. SIMILARITY: CONTRAINS 1 WD REPEAT (TRP-ASP DOMAIN).		
DR	EMBL; AB058693; BAB47419.1; InterPro; IPR000719; Euk_pkinase.		

ID TEST_RAT STANDARD: PRT: 628 AA.
 AC 063572:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Testis-specific protein kinase 1 (EC 2.7.1.-).
 GN TEST1
 OS Rattus norvegicus (rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Histar; TISSUE=Testis;
 RA MEDLINE=96125123; PubMed=8537404;
 RA Toshihisa J., Ohashi K., Okano I., Nunoue K., Kishikawa M., Kuma K.-I.,
 RA Miyata T., Hirai M., Baba T., Mizuno K.;
 RT Identification and characterization of a novel protein kinase,
 RT TEST1, specifically expressed in testicular germ cells.
 RL J. Biol. Chem. 270:31331-31337(1995).
 CC -1- FUNCTION: DISPLAYS SERINE/THREONINE-SPECIFIC PHOSPHORYLATION OF
 CC MYELIN BASIC PROTEIN (MBP) AND HISTONE IN VITRO. PROBABLY PLAYS A
 CC CENTRAL ROLE AT, AND AFTER THE MEIOTIC PHASE OF SPERMATOGENESIS..
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN TESTICULAR GERM
 CC CELLS.
 CC -1- DOMAIN: THE EXTRACATALYTIC C-TERMINAL PART IS HIGHLY RICH IN
 CC PROLINE RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: D50864; BAA09460.1;
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR004040; STR_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00221; STYK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 52 310 PROTEIN KINASE
 FT NP_BIND 58 66 ATP (BY SIMILARITY).
 FT BINDING 71 71 ATP (BY SIMILARITY).
 FT ACT_SITE 170 170 BY SIMILARITY.
 SQ SEQUENCE 628 AA; 67987 MW; F05P67BDD34B9AD CRC64;
 Alignment scores:
 Pred. No.: 3, 13e-06 Length: 628
 Score: 252.00 Matches: 133
 Percent Similarity: 35.56 Conservative: 71
 Best Local Similarity: 24.29 Mismatches: 272
 Query Match: 6.21% Indels: 134
 DB: 1 Gaps: 26
 US-09-836-392-8_COPY_22_2205 (1-2184) x TEST_RAT (1-628)
 QY 4 CTGAGCAGCTCGGGCCACCGATGCCATG-----AAGAAC 39
 DB 69 ValArgHisArgIleSerGlyIleValMetValLeuLysMetAsnLysLeuProSerAsn 88
 QY 40 TTCTCCGAGTTCGCGAGAGCCAGCATCGCTGACGACGCCCTGCATCGTG 99
 DB 89 ArgSerAsnThrLeuArgIleValGlnLeuMetAsnArgLeuArgHisProAsnIleLeu 108
 QY 100 GCGCTATGCGATCAACGATCCACCGCGCTGTGCTGCGCTGGAGAGCTGGCGGCTGAGC 159

DB 109 ArgPheMetGlyValCysValHisGln-----GlyGlnLeuHis 121
 QY 160 ACCCTCAACACCGCTGCTCGGAGAACCCAGAGAT-----TCTTCCTTTTATCCC 210
 DB 122 AlaLeuThrGlnLysMetAsnGlyThrLeuGlnLeuLeuSerSerProGluPro 141
 QY 211 CTGGGACATGCTCCACCAAAATAGCCTACAGATGCGCTGGCGCTGCTACCTG 270
 DB 142 LeuSerTrpProValArgLeuHisLeuAlaLeuAspIleAlaGlnGlyLeuArgIleLeu 161
 QY 271 CACAGAAAACATCATCTTCTGTGACCTGAGCGACACATTCCTGTGCTCCTT 330
 DB 162 HisAlaLysGlyValPheHisArgPheLeuThrSerLysAsnCysLeuValArgArgGlu 181
 QY 331 GACGTCAAGGACACATCAATCAAGCTATGTACTACAGGATTCGAGCGAG----- 384
 DB 182 Asp-----GlyGlyPheThrAlaValAlaGlyAspPheGlyLeuAlaLysIlePro 199
 QY 385 TCATTCATGAGGAGCGCC-----CTAGCGGTGGAGGCGACCTGCTGCTACCG 432
 DB 200 ValTyrArgGlnGlyAlaArgLysGluProLeuAlaValAlaGlySerProTyrTrpMet 219
 QY 433 GCGCCAGATCAAGCGCTGCGATTTGATATGATGAGAGAGATGATTTCTCTATGGA 492
 DB 220 AlaProGlnValLeuArgGlyGlnLeuLysIleValAlaAspValPheAlaPheGly 239
 QY 493 ATGGTCTCTACGAGTGTCTCA-----GGACAGCGCCCTGCACCTG 534
 DB 240 IleValLeuGlyGlnLeuLeuAlaArgValProAlaAspProAspPyrLeuProArgThr 259
 QY 535 GCGCACACCAAGCTCCAGATTTGCCAAGAGCTGCCAAGGCGATCGCGGTGTGGGG 594
 DB 260 GluAspPheGlyLeuAspAl-----ProAlaPheArgThrLeuValGly 274
 QY 595 CAGCCGAGAGAGATGAGTGTGCTGCGGCGACTGACGCGCTGATGAGAGTCTGGGACACT 654
 DB 275 AsnAspCysProLeuProHe-----LeuLeuAlaIleHisCysCysSerMet 291
 QY 655 AAGCCAGAGAGCGAGCGCTGCGCTGCGGTGCTGAGCGCAGATGAGAGACCGACTTTT 714
 DB 292 GluProSerAlaArgAlaProPheThrGlnIleThrGlnHisLeuGln 308
 QY 715 GGCACCTTATGATGATGACTGCTGCTGCGGAGACAGACCGCTTCTTCATCCAGGGC 774
 DB 309 -----IleLeuGlnGlnLeuProGluProThrProLeuAlaLysMetPro----- 323
 QY 775 CAGAGTACACCGTGTGTTGGATGAGAAAGAGAGCTCCAGAACTACACGCTGTG 834
 DB 324 -----LeuAlaLysAlaProLeuThrTyrAsnGlnIle----- 334
 QY 835 AAGCAGAGAGAGCGCTCATGAGAGGTGACAGAGAGTGTGCTGCCCTGGAGTGAAGTGACC 894
 DB 335 -----SerValProArgGlyGlyProSerAlaThrLeuProArgSerProArgL 352
 QY 895 TGCCA-----GCTCCAGGTCCAGATCCCTGTGGAC----- 926
 DB 352 euserArgSerArgSerArgPheLeuProSerProGluSerProProSerTrpG 372
 QY 927 -----AGCCACCGAGGACAGAAATATCATCATCATCATCATCATCATCATCATCAT 969
 DB 372 LysAspAsnLeuThrArgValAsnProPheSerLeuArgGluAspLeuArgGlyGlyLys 391
 QY 970 -----ATGTGCCCTTTAAACACACCCCAA 993
 DB 392 IleLysLeuLeuAspThrProCysLysProAlaThrProLeuProLeuValProProSer 411
 QY 994 CAGGCTTGATATCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1053
 DB 412 ProLeuThrSerThrGlnLeuProLeuValAlaSerProGluSerLeuValGlnProGlu 431
 QY 1054 AATTCCTACCTGCTTGTAGGCGGCTGCGGATGGGCTGTGCTGCTGCTGCTGCTGCTGCTG 1113

QY 1543 AGGCCAACCAAGGTGCTGAGGGGACACTCCATCGCGAGCTGAGCATCTGACACT 1602
 |||||
 DB 585 CARGPRTHTPRTAAlaValAlaAlaArgTyrArgAsn-----LeuAsnCysGluAl 601
 QY 1603 GAGGAGCTGGGACGACGATCTGATCCACGAGATCACTGACTACTGCTCCATG 1662
 |||||
 DB 601 aGlySerLeuLeu-----CysHisArgGlyHisHisAlaLys----- 613
 QY 1663 TCTCTACTCTCTATCCACCCGCGCTGCGAGCT 1702
 |||||
 DB 614 -ProProThrPro-----SerLeuGlnLeuProGly 623
 RESULT 14
 SRK4_SPOLA
 ID SRK4_SPOLA STANDARD: PRT: 506 AA.
 AC P42690.
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase SRK4 (EC 2.7.1.112).
 GN SRK4.
 OS Spongilla lacustris (Freshwater sponge).
 OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
 OC Haplosclerida; Spongillidae; Spongilla.
 OX NCBI_TaxID=6055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:92334672; PubMed:1378585;
 RA Oettille S., Raulf F., Barnekow A., Hannig G., Scharf M.;
 RT "Multiple src-related kinase genes, srk1-4, in the fresh water sponge
 Spongilla lacustris.";
 RL Oncogene 7:1625-1630(1992).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (AC P42686) and 4 (shown
 here); are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X61604; CAA43801.1; .
 DR HSSP: P00523; 2PTRK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00401; SH3DOMAIN.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD0000061; Euk_pkinase; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 1.
 DR SMART: SM00219; TYRKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00001; SH2; 1.
 DR PROSITE: PS00002; SH3; 1.
 DR Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;

KM SH3 domain; Phosphorylation; Alternative splicing.
 FT DOMAIN 54 116 SH3.
 FT DOMAIN 122 214 SH2.
 FT NP_BIND 246 254 ATP (BY SIMILARITY).
 FT BINDING 246 254 ATP (BY SIMILARITY).
 FT ACT_SITE 359 359 BY SIMILARITY.
 SQ SEQUENCE 506 AA; 57561 MW; 8ABDF4A2546C280B CRC64;
 Alignment Scores:
 Pred. No.: 2,29e-06 Length: 506
 Score: 254.50 Matches: 73
 Percent Similarity: 49.38% Conservative: 46
 Best Local Similarity: 30.29% Mismatches: 87
 Query Match: 6.27% Indels: 35
 DB: 1 Gaps: 8
 US-09-836-392-8_COPY.22.2205 (1-2184) x SRK4_SPOLA (1-506)
 QY 37 AACTCTCCGAGTTCGCGAGGAGCGACATGCTGACAGCGCTGACAGCCCTGCATC 96
 |||||
 DB 276 SERLLEGLUGLPHLeuGlnLeuGlnLysSerLeuGlnLeuArgHisProLysLeu 295
 QY 97 GGGCGCTCATGCGCATGACATC-----CACCGCTCTGCTGCGCTGAGCTCGCG 150
 |||||
 DB 296 LLeGlnLeuTyrAlaValAlaCysThrLysGlnGlnProLeuTyrLLeValThrGlnLeuMet 315
 QY 151 CCGCTCAGACGCTTCACACCGCTGCTCCGAGAACGCCAGATCTTCTTATACCC 210
 |||||
 DB 316 LysHisGlySerLeuGlnLeuArgLysArgGlySerLeuLysLeuProAsp 335
 QY 211 CTGGAGACACATGCTACCCCAAAATAAGCTTCACAGATGCGCTGCGGCTGCTACCTG 270
 |||||
 DB 336 LeuValAspMetCysSer-----GlnValAlaSerGlyMetSerTyrLeu 350
 QY 271 CACAGAAAACATCATCTTCTGTGACCTGAGCTGAGCGACACATCTGCTGCTCCTT 330
 |||||
 DB 351 GlnGlnGlnAsnTyrLLeHisArgAspLeuAlaValArgHisLLeu----- 366
 QY 331 GACGTCAAGAGGACATCAACATCAAGCTATGCTACAGGATTTCCAGG----- 381
 |||||
 DB 367 --ValGlyGlnHisLysLLeCysLysValAlaAspPheGlyLeuAlaArgValLLeAsp 385
 QY 382 CAGTATTCATGAGGAGCGCCCTAGCGGAGGCGACCTGCTGCTACAGGCCCCAGAG 441
 |||||
 DB 386 GlnGlnLLeTyrGlnAlaLysLeuLysPheProLLeLysTyrPheAlaProGln 405
 QY 442 ATCAGCGCTCGCATTTGTATGATGAGAGAGGTAGATATGCTCTCATGAGATGCTGCTC 501
 |||||
 DB 406 AlaAlaMetTyrSerArgPheThrLLeLysSerAspValTyrSerPheGlyLLeValLeu 425
 QY 502 TACGAGTGTCTGTC---GGACAGCGCCCTGACCTGGCGCCACACAGCTCAGATTGCC 558
 |||||
 DB 426 TyrGlnValLLeThrTyrGlnPheProTyrProGlyMetThrAsnAlaLLeValLeu 445
 QY 559 AAGAAGCTGTCCAGAGGATCGC-----CCGTTTCGGGAGCGGACCGAGAGATGTCAG 612
 |||||
 DB 446 GlnGlnLLeGlnGlnSerTyrArgMetProArgPrometGlyCysProGln----- 462
 QY 613 TTCCGCGAGCTCAGCGCGCTCATGATGAGTGTGCGACACTAAGCAGAGAGAGAGCGG 672
 |||||
 DB 463 -----LysLeuTyrAlaLLeMetMetLaspCysTyrArgGlnAspProLLeSerArg--- 479
 QY 673 CTGGCGCTGTGTGTGTGAGCCAGATGAAGGACCGGACTTTGCGACCTTCATGTATGAA 732
 |||||
 DB 480 -----ProThrPheGlnThrLeuSerTyrPhe 488
 QY 733 CTG 735
 |||||
 DB 489 Leu 489
 RESULT 15
 TSK_RAT

DR MM: 601782; Euk_Pkinase.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR004040; SRY_Pkinase.
DR InterPro: IPR001245; TYR_Pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Euk_Pkinase; 1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 57 314 PROTEIN_KINASE.
FT NP_BIND 63 71 ATP (BY SIMILARITY).
FT BINDING 86 86 ATP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 626 AA: 67712 MM; 3BBD96C20F29BEO CRC64.

Alignment Scores:
Pred. No.: 2,06e-06 Length: 626
Score: 255.50 Matches: 150
Percent Similarity: 33.69% Conservative: 71
Best Local Similarity: 22.87% Mismatches: 242
Query Match: 6.30% Indels: 193
DB: Gaps: 28

US-09-836-392-8_COPY_22_2205 (1-2184) x TEST_HUMAN (1-626).

QY 4 CTGAGGACCTCGGCGCCACGATGCGATG-----AAGAAC 39
DB 74 VALARGHLSARGLHNSERGLYGLNVALMETVALLEULYSMETASNLYSLEUPROSERASN 93
QY 40 TCTCCGAGTTCGCGGACGACGATGCGACGCGCTGCGACGACCCCTGCATCGTG 99
DB 94 ARGGLYASNTHLLEULARGGLULVALGlnleuMetasnlrlyleuArgHlsProasnIleleu 113
QY 100 GCGCTCATCGGACATGACATCCACCGCTCTGCTTCCGCGGACGCGCGCTCAC 159
DB 114 ARGPHMETGLYVALCYSLVALHLSGLN-----GLYGLNLEUHLIS 126
QY 160 AGCCTACACCGTGTCTGCTCCGAGAACGCCAGAGAT-----TCTTCCTTTARACC 210
DB 127 ALALEUTHGLULYMETASNLGLYTHLLEUGLULLEULEUSERPROGLUPTRO 146
QY 211 CCGGACACATGCTGACCCCAAAAATAGCTTACAGATCGCTCGGCGCTGCGCTCAT 270
DB 147 LEUSERTPROVALARGLEUHLISLEULALEULASPLLEALLARGGLYLEULARGTYRLEU 166
QY 271 CACNAAGAAAACATCATCTTCTGTGACTGAGTGCAGACATCTGTGTGTGCTCCTT 330
DB 167 HISSELYGLYALPHENHSARGASPLEUTHSERLYSASNLYSLEUVALARGARGGLU 186
QY 331 GACGTCAAGAGACATCAACATCAAGCTATCTGACTACGGGATTTGAGGCGAG----- 384
DB 187 ASP-----ARGGLYPHERHRLAVALGVALGVALSPHEGLYLEULAGLUYSILEPRO 204
QY 385 TCAITTCATGAGGCGCC-----CTAGGCGTGAGGCGCTCTGCTGCTGACAC 432
DB 205 VALTYRARGGLULGVALARGLYSGULPROLEUALAVALGVALYSERPROTYRTPMET 224
QY 433 GCCCAGAGATCAAGCGCTGCTATATATGATGAGAGAGATAGATATGTTCTCTATGA 492
DB 225 ALAPROGLUVALLEULARGGLYGLULLEUTYRASPGLULYSALASPALPHEALAPHEGLY 244
QY 493 ATGGTCTCTACGAGTTGCTGCA-----GGACAGCCCTGCACTG 534
DB 245 ILEVALLEULYSGLEULEULLEALARGVALPROALASPROMASYRLEUPROARGTYR 264
QY 535 GGGCACCACACAGTCCAGATGCGCAGAGAACCTGTCCAGGGCATCCCGCGGTCTGCGG 594
DB 265 GLUASPHHEGLYLEULASPAL-----PROALAPHEAETHLLEULVALGLY 279
QY 595 CACCCGAGAGATGCACTCCGCGCGACGTCAGGCGCTCATGATGAGTCTGCGACACT 654

DB 280 ASPASPCYSPROLEUPROPE-----LEULEULEUALLEHLSYSCYSASNLEU 296
QY 655 AAGCAGAGAAAGCGACCGCTGCGCTGCTG-----GTGCTGAGC 693
DB 297 GLUPROSERTHRARGALAPROPHETHRGLULETHRGLNHLISLEUGLUTRPLILEUGLU 316
QY 694 CAGATGAAGAACCCGACTTTTCCACCTTCATGATGACTGTGCTGTGGAGACAGACA 753
DB 317 GLNLEUPROGLUPROALAPROLEUTHR-----ARGTHR 327
QY 754 GCCTTCTTCATCCAGGCGCAGAGATACACCGTGTGTTGGAGATGAAGAAGAGAG 813
DB 328 ALALEUTHRHLSPASGLNGLY-----SERVALALARGLY----- 339
QY 814 TCCAGAACTACACCGTGTGTGACACAGAGAAGAGGCTCATGAGAGTGCAGAGATGTC 873
DB 340 -----GLYPRO----- 341
QY 874 TGCCCTGGGAGAGAGTGTGAGTCCAGCTCCAGGTCCAGAGATCCCTGTGCACGCCACC 933
DB 342 -----SERALATHLEUPROARGPROASPROARGLEU-----SER 354
QY 934 GAGGACAGAAATCTACATCTACACCTCA----- 965
DB 354 RGSERARGSERASPLEUPHEUPROPROSERPROGLUSERPROBASNTRPGLYASPA 374
QY 966 -----GGGCATGTGCC 978
DB 374 SNLEUTHRARGVALSNPROPHESERLEUARGGLUASPLEULARGGLYLSILEYSL 394
QY 979 TTAACACACCCCAACAGGCTTGATCTCAGTGTGCTGCTGCTGCTGCTGCTGCTG 1038
DB 394 EULENASPHRPROSER-----LYSPROVALPROLEUVALPROPROSERPHROPISE 413
QY 1039 CCTGTATTAATAAGATCTACCTGCTTACGAGGCGCTGCGCGATGGCTGTGCTGCT 1098
DB 413 RTHRGLNLEUPROLEUVALTHRTHRPROGLULTHRLEUVALGLNPRO----- 428
QY 1099 GTGTTCCCGTGTGCGGCGACCCCAAG-----ACAGCTGCTCTAC 1143
DB 429 -----GLYTHRPROALARGARGCYASARGSERLEUPROSERPROI 443
QY 1144 CTGTGCTCACACACGCCACA-----GCTCCAG 1173
DB 443 ULEUPROARGARGMETGLUTHRALALEUPROGLYPROGLYPROALVALGLYPROSE 463
QY 1174 TTCAGATGCGCGATTAAGACGACGACGAGAACCCCTTACCAGTGAAGCCATGAGG 1233
DB 463 RALAGLUGLULYSMETGLUCYSGULYSERPROGLUPROGLUPROGLYPROAL 483
QY 1234 GTCAACAGGCGCTCTGAGGCTGTGACAGCAATGGCGCGGCTCTGTCATGACACTGT 1293
DB 483 APROGLNLEUPROLEUALAVALATHRASPNPHEIUSERTHCYSSERSERLASE 503
QY 1294 GCCTCCCTG----- 1303
DB 503 RGLNPROTRIPSERPROARGSERGLYPROVALLEUASNAASNPROPROALVALAVALA 523
QY 1304 -AGATCTGACGCGCTGAGACGCTTACATGCGCCCTCATGTTACGTACGACGTGTC 1362
DB 523 IASNSERPROGLNGLYTRPALA-----GLYGLUPROTPASNAARGALINHSSE 540
QY 1363 AGCTTGAGGCGAGAGGAGGAGAGTGTGCTGCTGATGACAGAACCACTCTTG 1422
DB 540 RLEUPROARGALA-----ALALALEUGLARGTHLEUPRO----- 552
QY 1423 GTGATGTACACTCCACACCTACACAGCTGTGTGCCCGGTACTTCTGCGGGTCCAC 1482
DB 553 -----SERPROBROSERALAPROARGGLUPROASPLUCLYLEUPROCY 568
QY 1483 CCCCTCAGGACATGTTCCCGTGGCGGCTTGACACAGGAAACCCCGGACGACACAC 1542
DB 568 SPRO-----GLYCYSGLYSEU-----GLYPROPHESERPHHEGLYPHELEUSERMETCYSPT 585

```

XN NCBI_TaxID=44689;
RX [.]
RP SEQUENCE FROM N.A.
RA MEDLINE=90287147; PubMed=1972546;
RT Tan J.L., Spudich J.A.;
RL "Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoideum"; Mol. Cell. Biol. 10:3578-3583(1990).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
-----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-slb.ch/announce/ or send an email to license@isb-slb.ch)
-----
CC EMBL: M33784; AAA33203.1; .
CC PIR: B35670; B35670.
CC HSSP: P08631; IAD5.
CC DictYdb: DD03011; PYKB.
CC InterPro: IPRO00719; Euk_Pkinase.
CC InterPro: IPRO00400; Sty_Pkinase.
CC InterPro: IPRO01245; Tyr_Pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_Pkinase; 1.
CC SMART: SMO0221; STYKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT NON_TER 1 1
FT DOMAIN 108 381 PROTEIN KINASE.
FT NP_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E93918B605BAEC1 CRC64;

Alignment Scores:
Pred. No.: 1.67e-06 Length: 410
Score: 257.00 Matches: 76
Percent Similarity: 45.26% Conservative: 48
Best Local Similarity: 27.74% Mismatches: 114
Query Match: 6.33% Indels: 36
Gaps: 9

US-09-836-392-8_COPY-22_2205 (1-12184) x RYK2.DICDI (1-410)
QY 4 CTGAGCAGCACTGGCGGCACACGATGCCATGAAGAATTCTCCGAG---TTCCGCGAGGAG 60
Db :||||| :|||: ||| ||| ||| :|||
QY 134 Ilelyslyslslleuylilielelglyslpeluglnphelysglinrghellearglu 153
Db :||||| :|||: ||| ||| ||| :|||
QY 61 GCCAGCATGCTGCACGCGCTGCAGCACCCCTGCATCGTGCGCGTCATGGCGCATCAGCATC 120
Db :||||| :|||: ||| ||| ||| :|||
QY 154 Valglnasnlleuyllylglyasnhsnglnsnllevalmerphellleulylacystyr 173
Db :||||| :|||: ||| ||| ||| :|||
QY 121 CACCGGCTGTGGTTCGGCCCTGAGAGCTCGCGCGCTCAGACGAGCTTAACACCCTGCTGTCC 180
Db :||||| :|||: ||| ||| ||| :|||
QY 174 Lysprralacyslllelrnthrciluyrmetalaglylsrleutryasnilleuhs 193
Db :||||| :|||: ||| ||| ||| :|||
QY 181 GAGAACGCCAGATGATCTTCCTTTATACCCCCTGGGAGACATGCTCACCOCAAAAATAAGCC 240
Db :||||| :|||: ||| ||| ||| :|||
QY 194 Asnprdansestethrprolysalylstyrserpheprolevalleuylshetala 213
Db :||||| :|||: ||| ||| ||| :|||
QY 241 TACCAAGATCGCTCGGCGCTGGCTACTGTCACAAGAAAAACATCATCTTCTGTGACCTG 300
Db :||||| :|||: ||| ||| ||| :|||
QY 214 Thrspmetatalenuglyleuendhlisleuhsserlllethrlllevalhsargspneu 233
Db :||||| :|||: ||| ||| ||| :|||
QY 301 AANGTGGACAAACATTCGTGTGTGTGTCCTTACAGTCGAAGGACACATCAATCAAGCTA 360
Db :||||| :|||: ||| ||| ||| :|||

```

Dn 234 Thrsrclnshnlleleuleuaspdluendly-----Asnilleysile 248

Qy 361 TCGACTACAGGGGATTTTCAGGCAGTCAATCCATGAGGCCCTCAGCGTAGAGGC-- 417
|||||:|||||: ||| :|||:|||||: |||
Db 249 SerasppheciyleuseralaglulysSeratrglucylsermethtermsngly 268
418 -----ACTCCTGGCTACCAGGCCCCANAGATCAGGCTCGATT--GTATNGAT 465
|||||: ||| :|||||: ||| :||| :||| :|||
Db 269 GlylleCyasaProlaygtrparprofoluLeuthrlYsaSnleuglyHlsIysr 288
466 GAGAAGGTAATATGTTCTCATGAGATGGTGCTCAAGATTGCTCAGAGCACGCC 525
|||||:|||||: ||| :|||:|||||: ||| :|||:|||||: |||
Db 289 GlutysValaspvaltyrcyspheserleuValaitrgluilleLeuthrlcylsulle 308
Qy 526 CCGTCGACTG-----GGCCACCACCAAGCTCCAGATTGCCAMAGAACTGTCCAG 573
||| :||| :||| :||| :||| :||| :|||
Db 309 PropheresAspleuaspglySerclnargseralaglnAlala-----Tyrala 325
Qy 574 GGCAATCGCCCCGTTTCGTGGGAG-----CGGAGGAAGTCAATCCGGGAGATG 624
|||||:|||||: ||| :||| :||| :||| :||| :|||
Db 326 GlyleuatrgProfoleleProglutrycysaspproglu-----Leu 339
Qy 625 CAGGCGCATGATGAGATGGATGGGACACTAAGCCAGAGAACGACCGCTGGCCCTGCG 684
::: |||||: ||| :||| :||| :||| :||| :|||
Db 340 LysleuleuleuthrlglnCysrtlpoluilaaspprobasapargpropropheThrtyr 359
Qy 685 GTGGTGACCCAGATGAAGAGACCCGACTTTTCCACCTTCATGATGAAGACTGTGCTGGG 744
:::|||||: ||| :||| :||| :||| :||| :|||
Db 360 lIeValasnlysleulyt-----Glullesertprasn 370
Qy 745 AAGCAGACAGCCCTTCTCTCATCCAGGCGCAGAGATCAACC 786
Db 371 AsnProllleglyphevalSeraspInphenetyrglnTyrsr 384

RESULT 13

TEST_K_HUMAN STANDARD; PRT; 626 AA.

AC Q15569;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Testis-specific protein kinase I (EC 2.7.1.-).
GN TESTK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hepatoma;
RX MEDLINE=96125123; PubMed=8537404;
RA Yoshima J., Ohashi K., Okano I., Nunoue K., Kishloka M., Kunu K.-I.,
Miyata T., Hirai M., Baba T., Mizuno K.,
RT "Identification and characterization of a novel protein kinase,
TESTK1, specifically expressed in testicular germ cells.";
RL J. Biol. Chem. 270:31331-31337(1995).
CC -1- FUNCTION: DISPLAYS SERINE/THEONINE-SPECIFIC PHOSPHORYLATION OF
MEKLIN BASIC PROTEIN (MBP) AND HISTONE IN VITRO. PROBABLY PLAYS A
CENTRAL ROLE AT AND AFTER THE MEIOTIC PHASE OF SPERMATOGENESIS.
CC -1- DOMAIN: THE EXTRACTALYTIC C-TERMINAL PART IS HIGHLY RICH IN
PROLIN RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

EMBL; D50863; BAA09459.1; .
Genew; HGNC:11731; TESTK1.

```

Db 189 AAlaProGluValIleMetSerGlnHisTyrAspGlyLysAlaAspLeuPsrIleGly 208
OY 493 ATGGCTCTTACAGAGTGTCTGTACAGACAGCCCT-----GCATGGGCGCCACACAG 546
Db 209 ThrIleValIleGlnCysLeuThrGlyLysAlaProPheGlnIleAspSerProGlnAsp 228
OY 547 CTCACAGATTCCCAAGAGCTGTCCAAAGGAGCCCGGCTTGGGGACCGGAGAA 606
Db 229 LeuArgLeuPheThrGlyLysAsnLysThrLeuValProThrIleProArgGluThrSer 248
OY 607 GTCGCAATTCGGGAGCTGCAGGCGCTCATGATGAG- 642
Db 249 AlaProLeuArgIleLeuLeuLeuAlaLeuGlnArgAsnHisLysAspArgMetAsp 268
OY 643 -----TGTGGAGACTAAGCCAGAGAGAGAGCCGCTGCC 678
Db 269 PheAspGluPhePheHisThrPheLeuAspAlaSerProSerValArg----- 285
OY 679 CTCGCGGTGTGAGCCAGATGAAGAGCCGACTTTGCCA---CCTTCATGATGAATG 735
Db 286 -----LysSerProPro-ValProValProSerTyrProSer 298
OY 736 TGTGTGGAGAGCAGACCTTCTCTCATCCAGGCGCAGAGATACACCGTGGTGT 795
Db 298 rGlySerGlySerSerSerSerSerSerThrSerHisLeuAlaSerProSer 318
OY 796 TGGG-----ATGGAAG-----AGGATCCAGAAACACAGGTGT- 834
Db 318 uGlyGluMetGlnGlnLeuGlnLysThrLeuAlaSerProAlaAspThr-AlaGlyPheL 338
OY 835 AACACAGAAAGGCTCATGAGGTGCAGAGAT----- 869
Db 338 euHisSerSerArgAspSerGlyLysThrLysAspSerSerCysAspThrAspAspPhe 358
OY 870 --GTGCTGCCCTGGGAT---GAAGTGAGCTG---CCAGCTCCAGGCGCAGACATC 918
Db 358 AlMetValProAlaGlnPheProGlyAspLeuValAlaGluAlaProSerAlaLysProp 378
OY 919 CTGTGACAGC----- 948
Db 378 rPrAspSerLeuMetCysSerGlySerSerLeuValAlaSerAlaGlyLeuGlnSerH 398
OY 949 T-----ACATCTACACCTCAAGGCAATGTCCCTTAACACACCCACAGGCTT 1001
Db 398 IsGlyArgThrProSerProSerProGlySer----- 409
OY 1002 GATACTCCAGCTGTGTACCTGCTCTTGGCCGTGCTGTTATTAAGAAATTCCTA 1061
Db 410 -----SerSerProSerPro- 414
OY 1062 CCTGTCTTAGCGGCGCTGCGGATGGGCTGTGTGTTCCCGTGGTGGGCGAC 1121
Db 415 -----SerGlyArgAlaGlyProPheSer-----SerSerArgCysGlyAla- 428
OY 1122 CCCAAAGACAGCTGCTCTTACCTGTCTACACACAGCAACAGGTCCCAAGTTCAGAT 1181
Db 429 -----SerValProIle-----ProValProThrGlnValGlnAsnTyrG 442
OY 1182 CCGGATGAAGAGCGAGCGGAGAACCTACCCAGTAGAGGATGAGGTGTCTAAGC 1241
Db 442 InArgIleGlnArgAsnLeuGlnSerProThrGlnPheGlnThrProArg---SerSer 461
OY 1242 CGGCTGTAGTCTGTATACAGA-----ATGGCGCGGCTCTGTCTATCATCATGTCG 1295
Db 461 lAlleArgArgSerGlySerThrSerProLeuGlyPheAlaArgAlaSerProSerPro 481
OY 1296 CTCCTCGAGATCTGCAGCGGCTGAGCCCTACATAGGCCCTCATGGTATGTCAGT 1355
Db 481 rAlaHisAlaGlnHisGlyGly----- 488
OY 1356 CGTGTACACTCTGAGGAGGAGGAGAGAGTGTGCTGTGCTGGATGACAGGCCAA 1415

```

```

Db 489 -----ValLeuAlaArgLysMetSerLeuGlyGlyArgProTyrThrPro 505
OY 1416 CTCCTTGTGTAGTATCCACTCCACCACTACAGAGCTGTGTGCCGCTACTTGTGGGGGT 1475
Db 505 ePrProLValGlyThrIleProGluArg-----ProGlyTyrPsr---GlyT 520
OY 1476 CCCCAGCCCCCTCAGAGACATGTTCCGTGGGCGCTTGGACAGGAAACCCCGGACG 1535
Db 520 hPrProSerProGlnGlyAlaGluMetArgGly-----ArgSerProArgProG 537
OY 1536 CAGCCACGGGCCAACCCAAAGGTGCTGAGGGGAGCTCCATCGGAGAGTATCAT 1595
Db 537 LysSerAlaProGluHis-----SerProArgThrSerGlyLeuGlyC 552
OY 1596 GTACATGAGAGAGCTGGGAGCAGAGATCCGTATCCACAGAGATCACTACTGACTAGT 1655
Db 552 yAsArgLeuHisSerAlaProAsnLeuSer-----AspLeuHisValVala 567
OY 1656 CTCACATGCTCTCTACTCTCATCCACCCG---GCCAGGCTGCCAGGTCCCGCTCAG 1712
Db 567 rPrProLysLeuProLysProProThrAspProLeuGlyAlaValPheSerProProGln 587
OY 1713 CCTCCCCA----- 1720
Db 587 lAsrProProGlnProSerHisGlyLeuGlnSerCysArgAsnLeuArgGlySerProL 607
OY 1721 -----GCTCCCGACAG 1733
Db 607 yLeuProAspPheLeuGlnArgAsnProLeuProProGlyLeuGlySerProThrLysA 627
OY 1734 TTTCTCAATGTGCTTTTCCACGAGCTGCGAGACTAGACATGCTCAATAGCCCGG 1793
Db 627 lAlaLProSerPheAspPheProLysThrProSerSerGlnAsn-----LeuLeuAlaL 645
OY 1794 TGTGCTCCGACAGAGCTGTAGACCTGAGCC----- 1828
Db 645 euLeuAlaArgGlnGlyValAlaMetThrProProArgAsnArgThrLeuProAspLeu 665
OY 1829 -----CCATGAGCGGAGAGACTTCAAGCCAGCTGACGCGGTGAGAG- 1873
Db 665 eGluValGlyProPheHisGlyGlnProLeuGlyProGlyLeuArgThrGlyLysAsp 685
OY 1874 --TCTCGCCGTCAGAGACTATTTGGTCCCGGAGCGGCTGGAGATGTATGCTCAT 1931
Db 685 rGlyGlyProPheGlyArgSerPheSerThrSerArgLeuThrAspLeuLeuLysA 705
OY 1932 TGGCGCTG---AGAGAGATTCTGAAGCCAGCGGCGCGAGCATGTCTTAAGC 1988
Db 705 lAlaLProGlyThrGlnAlaProAspProGlySerThrGlnSerLeuGlnLysProX 725
OY 1989 CCGAGAGCTGACTCCCATGAGGCTGCTGGTGTGATGCTG---CCGTTGTGCAAGAGAC 2045
Db 725 eGluIleAlaProSerAlaGlyPheGlyGlySerLeuHisProGlyAlaArgAlaGlyG 745
OY 2046 TGTGTGTGACCTTTGAAATGAATAACACAGAGTGTGCTGGCGCTGGAGGGGCTG 2105
Db 745 lYThrSerSerProSerProValAlaPheThrValGlySerProProSerGlySerThr 765
OY 2106 GGGCGCCAGGAGGTGACA-----TTTCTACAGCTGCTTCCAGAGAGC 2149
Db 765 rPrProGlnLysProArgThrArgMetPheSerAlaGlyProThrGlySer 781

```

```

RESULT 12
ID KYK2.DICDI STANDARD; PRT; 410 AA.
AC P18161;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).
GN PYK2 OR DPYK2.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostellium.

```


Oy	2170	TGCAGGCTCCAGCGGGCGCA-----GCTCCTCGAGACGTGGTGAAGAATGTCGAACTCCC	2114
Dd	19	CysProPdpProCysLeuProGlyAlaProAspGlnThrArgTrpLeuProProGly	38
Oy	2113	TGG-----GGCCCCAGCCCTCCACA-----	20933
Dd	39	TrpGlyInnArgThrAlaProThnGlnValGlyLeuAlaAspAlaAlaSerProAspLu	58
Oy	2093	-----	20933
Dd	59	LeuGlnAspGlnAlaSerGlyAlaArgProGlyGlyYasnArgValGlyAlaGlyArg	78
Oy	2092	---GGCCAGGCCACCACACTCTGTCTTTTCATTTCGAAGGTGCACACACAGTCTCTTG	20366
Dd	79	GlyAraProGlyThr-----	85
Oy	2035	CCACGACGGCGCAT-----CCACGACGCCCATGGGAGACGCTCGGGCTT	19858
Dd	86	ProSerArgInserAlaGlyThnGlyProAlaGluGlnAlaAspIslaInHisSer---	104
Oy	1984	TTAAGACGGCAATGACTCGGGCCSCCGCTTCAGAAATCTTTCTCCAGGCCCAATGACGA	19235
Dd	105	-----AsnProThnGly-----	108
Oy	1924	TAACTCTCCACCGCGCTGGGACCCAAATGAGTCTGTGACGGCGAGAACTTCACGG	18655
Dd	109	-----GlyCysSerasp	112
Oy	1864	CCT-----	18566
Dd	113	ProGlnAlaSerProAlaGlyThrArgAlaInnAlaGlyTYrAlaLeuGlyLysLysSerIleActy	132
Oy	1855	GCTGGCTGAAGTGTCCCGCTCATGGGGGTCAAGTCAATGCTCACAGCTGTGGAGCGAG	17968
Dd	133	LeuGlySerArgGlyProArgTrpOnnAsProAlaRheGlnValaGlnIlnPrSerAlaArgAsn	152
Oy	1795	CACCGGGGGTATGNASATGTCTGAGTCTCGSAGTCSGGTGGAGAAAGGACACASTGGAG	17363
Dd	153	-----ProGlyCys-----ProArgThr-----TriParArgArgSer-----	163

QY	1735	AACTGCTGGGAGCTGGGAGGCTTGAAGGGGAGACCTGGACCTGGGGG-----	16853
Db	164	-----GlyAlaGlnArgGlyHisProProProGlyAlaGlyGlnArg	177
QY	1684	-----GTGGGATGAGCACTAGAGACATGGAGCAGTAGTCAGTACTGATT	1637
Db	178	ProSerGlyProThrGlyArgProAlaAlaProGlyAlaProGlyThProAlaAla	197
QY	1636	CTTGCT--GGATCAGGATCTGGGTCCAGCT-----CTTCACCTGTACATGAGC	1589
Db	198	ProGlyProGlyGlyGlyAlaAlaAlaProSerGlyAlaThrProHisProGluArgGly	217
QY	1588	TCACGTCGGGATGAGTAGAGTCCGCTCAGGACACCTTTGGTGGCCGTGGCTGGCCG	1529
Db	218	SerGlyProAla---AsProProAlaAlaAlaAlaArgLeuProProGluArgGlnIuPro	236
QY	1528	GGGTCCTCGTGTCCAAAGGCCGCGAGGAAACATGCTCCCTGAGGGGGCTGGGACCCGC	1468
Db	237	ArgLeuProGlnAspLeuAlaAlaAlaGlnArgCysPro--AlaGlyProProProThr	255
QY	1468	AGAGTACCGGGCAGACACAGCTGGTAGGTGGAGTGGATCAACCAAGAGTTGGCT	1409
Db	256	ArgSer--GlyAlaAlaAlaGlnArg-----	263
QY	1408	TGTCATCCAGGACAGACAGACCTCTCCCTCCCTGCTCAGAGC-----	1364
Db	264	-----ThrHisArgArgProProGlyCysProArgSerAlaArgsnProGly	279
QY	1363	TGCACAGCACTAGCATACCATGAGG-----	1337
Db	280	CysProArg-----ThrTyrArgArgSerGlyAlaGlnArgGlyHisProPro	296
QY	1336	-----GGGCCATGAGGGCT--CCAGCCGCTGCAGACT	1304
Db	297	ProGlyAlaGlyGlnArgProSerGlyProThrGlyGlyArgProAlaAlaProGlyAla	316
QY	1303	CCAGGAGGACACACTGCATGACAAAGAGGCCGCGCCATGCTGTACACAGACCTCAGAC	1244
Db	317	ProGlyThrProAla-----AlaProGlyProGlyGlyGlyAlaAlaAlaProSerGly	334
QY	1243	CGCTGTGACACACCTGCATGCGCTTCACGTGGTAGGGGT-----TCTGCC	1199
Db	335	-----AlaThrProHisProGluArgGlySerGlyProAlaAsProProAlaAla	351
QY	1198	GTGCGTCTTCATCCGCGCATGCTGAACTTGGACCTGTGGCTGTGTGACACAGTAGG	1135
Db	352	AlaArgLeuProProGlu-----Arg	358
QY	1138	AGCAGC-----TGTCCTTGGGGTGCCCC	111
Db	359	GlnGluProArgLeuProGlnAspLeuAlaAlaGlnArgSerProAlaGlyProPro	378
QY	1114	GCACACGAGGAA-----ACACAGCCACAGGCCATGGCGAGGCCG-----CTAGA	1067
Db	379	ProThrArgSerGlyAlaAlaAlaAlaGlnArgThrHisArgArgProProGlyCysProArg	398
QY	1066	CCAGTAGGAATCTTTTAATACAGGACGCGCCACAGACGAGTGCAGACAGCTGGAG	1007
Db	399	SerAlaArgsn-----ProGlyCysProArgThr-----	408
QY	1006	TATCCAAAGGCTGTGGGGGTGTGTTAAAGGGGACATGCCCTTGAGGGGTGATAGTAGA	947
Db	409	TTrpArgArgArgSerGly---AlaGlnArgGlyHisProProGlyAlaGlyGlnArg	427
QY	946	TTTTGTGCTCTCGTGGCTGTCCACAGG-----ATCTGTGACCTGGACCTGGACAGC	893
Db	428	ProSerGlyPro-----ThrGlyGlyArgProAlaAlaProGlyAlaProGly	443
QY	892	TCACCTTCATCCAGGACAGACATCTCTGCA-----	860
Db	444	ThrProAlaAlaProGlyProGlyGlyGlyAlaAlaAlaAlaProSerGlyAlaThrProHis	463

```

Db 160 ArgArg-----Argser 163
QY 339 GGACACATCAACATCAAGCTATCTGACTAGGGATTTCAGAGGAGCTATTCATGAGG 398
Db 164 GYAlaGlnArg-----Gly 168
QY 399 CGCCCTAGCGCTGAGGGCACTCTGGCTACAGGCCAGAGATCAAGCCCTGCATTTGT 458
Db 169 HisProProGlyAla-----GlyGlnArgPro 178
QY 459 ATATGATGAGAGAGTAGTATGTTCTCTATGATGAGTGGCTCTACAGATTGCTGTACG 518
Db 178 ----- 178
QY 519 ACAGGCCCTGCTGCTGAGGCGACACAGCTCCAGATTGCCAAGAGCTGCCAGGCGAT 578
Db 179 SerGlyProHisGlyGlyArgProAlaAlaPro-----GlyAlaProGlyThr 194
QY 579 CCGCCCGCTTCTGGGCGAGCGGAGAGATGCTCCGGAGCTGCGAGCGCTCATGAT 638
Db 195 ProAlaAlaProGlyProGlyGlyAlaAlaValProSerGlyAlaThrProHisPro 214
QY 639 GGAATGCTGGGACACTAAGCCAGAGAACCCGCTGGCCCTGTGGCTGTGAGCCAGAT 698
Db 215 --GluArgGly--SerGlyProAlaAspProProAlaAlaAlaArgLeuProGlyUar 233
QY 699 GAAGACCCGACTTTTGGCACTTCATGATGATGACTGTGCTGGAGAGACAGACAGCT 758
Db 233 GGIuGluProArg--LeuPro-----GlnAspLeuAlaAlaGlnArgGlySerPro 250
QY 759 CTCTGTATCCAGGGCGCAGAGAT-----ACACCGT 788
Db 250 IagGlyProProProThrArgSerGlyAlaAlaAlaGlnArgThrHisArgArgProG 270
QY 789 GGTTTGTGGATGAGAAAGAGAGTCCAGAACTACAGCGTGTGAAACAGAGAGAGG 848
Db 270 IyCys--ProArgSerAlaArgAsnProGlyCysProArg--ThrTrpArgGly 287
QY 849 CCTATGAGAGGTGAGAGATGTCTGCGCTGGATGAGATGAGTGAAGTCCAGTCCAGT 908
Db 287 rgsSerGlyAlaGlnArgGlyHisProProGly-----AlaGlyGlnArgP 303
QY 909 CCAAGATCCCTGTGGAGAGCCAGCCAGAGACAGAAATCTACATCTACACCTCAAG- 967
Db 303 roSerGlyProThrGlyGlyArgProAlaAlaProGlyAlaProGlyThrProAlaAlaP 323
QY 968 -----GCATGTGCCCT-----TAAACACACCCCAACAGGCTT 1001
Db 323 roGlyProGlyGlyAlaAlaValProSerGlyAlaThrProHisProGlyUarGly- 342
QY 1002 GGATATCTCACAGCTGTCAGCTGCTTCTTGGCGTCTTTTAAAGAATTCCTTA 1061
Db 343 -----SerGlyProAlaAspProProAla-----AlaAlaArgLeuProP 356
QY 1062 CCTGTCTTACGGGCGCTCCGAGTGGCTGTGGCTGTCTTCCCGTGGCGGCGAC 1121
Db 356 roGlu-----ArgGlnGluProArg-----LeuP 364
QY 1122 CCCAAGAGACAGCTGCTCTCTGCTGTCTACACAGCCACAGAGTCCA----- 1171
Db 364 roGlnAspLeuAlaAla-----AlaGlnArgCysProAlaGlyProProProThr 381
QY 1172 --AGTTGAGCATCGCGGATGAGAGACGAGAGACCCCTACCCAGAGGAGGATGGA 1229
Db 381 rgsrGlyAlaAlaAlaGlnArgThrHisArgArgProPro----- 394
QY 1230 GGTGTCAACAGCGGCTCGAGGTCTGGTACAGCAATGGCGGCGCTCTTCATCGA 1289
Db 395 -----GlyCysProArgSerAlaArgAsnProG 404
QY 1290 CTGTGCT--CCTTGAGATCTGACAGCGGCTGAGAGCCCTACATGCGCCCTCCATNGT 1346

```

```

Db 404 IyCysProArgThrTrpArgArgSerGlyAlaGlnArgGlyHisProPro----- 422
QY 1347 TAGCTAGTGTGTGACAGCTGTGAGGGCAGAGGGAGAGGTCTGTGCTGGATGA 1406
Db 423 -----GlyAlaGlyGlnArgProSerGlyProThrGlyG 434
QY 1407 CAAGCCAACTCTTGGTATGATGATACCACTCCAGCAGCTACAGCTGTGTGCTT 1466
Db 434 IyArgProAla-----AlaProGlyAlaProGlyThrProAlaAlaProGlyProG 451
QY 1467 CTGGGGGTCCCGAGCCCTCAGAG--ACATGTTCCG----- 1504
Db 451 IyGlyGlyAlaAlaValProSerGlyAlaThrProHisProGlyUarGlySerGlyPro 471
QY 1505 -----TGGCGCTTGGACAGGAGACCCCGAGCGAGCC- 1540
Db 471 IAspProProAlaAlaAlaAlaArgLeuProProGlyUarGlyGlnArgLeuProGln 491
QY 1541 --ACAGGCGCAACCCAAAGTGTGCTGAGGGG--ACTCCATCGCGAGCTGACATCAT 1595
Db 491 spleuAlaAlaAlaGlnArgCysProAlaGlyProProProThrArgSerGlyAlaAla 511
QY 1596 GT-----ACAG 1601
Db 511 IagInArgThrHisArgArgProProGlyCysProArgSerAlaArgAsnProGlyCysP 531
QY 1602 TGAGAGCTGGCGAGCAGAGATCCCTGATCCAGCATCTACTACTACTCTCT- 1657
Db 531 roArgThrTrpArgArgArgSerGlyAlaGlnArgGlyHisProProProGlyUarGlyG 551
QY 1658 --CCTATCTCTCTACT--CCTCATCCCGACCCCGAGCTGCGAGTCCAGTCCCTC 1709
Db 551 IArgProSerGlyProThrGlyGlyArgProAlaAlaProGlyAlaProGlyThrPro- 570
QY 1710 AAGCTCCCGAGCTCCCGAGAAATCTTCCAGTGTGCTTCTCCAGCTCCAGAGA 1769
Db 571 --AlaAlaProGlyProGlyGlyAlaAlaAlaVal----- 581
QY 1770 CTCAGACATGCTATACAGCGCGGCTGCTGCTCCAGAGCTGAGCATGACGTGACCCC 1829
Db 582 -----ProSerGlyAlaThrPro-HisProGlu 590
QY 1830 CATGACGGGAGACCTTCAAGCCAGCAGCTGACAGCGGTGAAGATCTCGCCGTACAGA 1889
Db 591 ArgGlySerGlyProAlaAspProProAlaAlaAlaArgLeuProProGlnArgGlnGlu 610
QY 1890 CCTCATTTGGGTCCCGAG-----GCGCGGTGAGATGTTATCGTCATTGGCTTGA 1940
Db 611 ProArgLeu--ProGlnAspLeuAlaAlaGlnArgCys----- 623
QY 1941 GAAGATTTCTGAAGCCGAGCGGCGAGTGCATGCCGTCTTAAAGCCGAGAGCTGAC 2000
Db 624 -----ProAlaGlyProProProThrArgSerGlyAlaAlaGlnArg 638
QY 2001 TCCGATGGGTGCTGTGATGATGCTGCGCTGAGTGGAGAGACAGACTGTGT 2051
Db 639 ThrHisArgArgProProGlyCysProArgSerAlaArgAsnProGlyCys 655

```

```

RESULT 10
ID YH1_EBV STANDARD; PRT; 660 AA.
AC P03181;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical BHLFI protein.
OS Epstein-Barr virus (strain 95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;

```


GN ABL1 OR ABL OR JTK7.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic;
 RX MEDLINE=90082420; PubMed=2687768;
 RA Fainstein E., Elmat M., Gokkel E., Marcelle C., Croce C.M.,
 RA Gale R.P., Canaan E.;
 RT "Nucleotide sequence analysis of human abl and bcr-abl cDNAs";
 RL Oncogene 4:1477-1481(1989).
 RN [2]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=87028219; PubMed=3021337;
 RA Shivelman E., Lifshitz B., Gale R.P., Roe B.A., Canaan E.;
 RT "Alternative splicing of RNAs transcribed from the human abl gene and
 RT from the bcr-abl fused gene";
 RL Cell 47:277-284(1986).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS IA AND IB).
 RC TISSUE=Lung carcinoma;
 RX MEDLINE=95394474; PubMed=7665185;
 RA Chisoe S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,
 RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jan L., Ma Y.,
 RA Maury H.-J., Pan H.-Q., Sarhan O.H., Toth S., Wang Z., Zhang G.,
 RA Heisterkamp N., Groffen J., Roe B.A.;
 RT "Sequence and analysis of the human ABL gene, the BCR gene, and
 RT regions involved in the Philadelphia chromosomal translocation";
 RL Genomics 27:57-82(1995).
 RN [4]
 RP SEQUENCE OF 360-426 FROM N.A.
 RX MEDLINE=83245023; PubMed=6191223;
 RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;
 RT "Homology between phosphotyrosine acceptor site of human c-abl and
 RT viral oncogene products";
 RL Nature 304:167-169(1983).
 RN [5]
 RP SEQUENCE OF 27-40 FROM N.A.
 RX MEDLINE=88065859; PubMed=2825022;
 RA Fainstein E., Marcelle C., Rosner A., Canaan E., Gale R.P.,
 RA Drezner O., Smith S.D., Croce C.M.;
 RT "A new fused transcript in Philadelphia chromosome positive acute
 RT lymphocytic leukemia";
 RL Nature 330:386-388(1987).
 RN [6]
 RP SEQUENCE OF 825-845 FROM N.A.
 RX MEDLINE=94142331; PubMed=7545908;
 RA Inokuchi K., Futaki M., Dan K., Nomura T.;
 RT "Sequence analysis of the mutation at codon 834 and the sequence
 RT variation of codon 837 of c-abl gene";
 RL Leukemia 8:343-344(1994).
 RN [7]
 RP STRUCTURE BY NMR OF SH2 DOMAIN.
 RX MEDLINE=92370689; PubMed=1505033;
 RA Overduin M., Rios C.B., Mayer B.J., Baltimore D.;
 RT "Three-dimensional solution structure of the src homology 2 domain of
 RT c-abl";
 RL Cell 70:697-704(1992).
 RN [8]
 RP STRUCTURE BY NMR OF SH2 DOMAIN.
 RX MEDLINE=93101588; PubMed=1281542;
 RA Overduin M., Mayer B.J., Rios C.B., Baltimore D., Cowburn D.;
 RT "Secondary structure of Src homology 2 domain of c-Abl by
 RT heteronuclear NMR spectroscopy in solution";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11673-11677(1992).
 RN [9]
 RP STRUCTURE BY NMR OF SH3 DOMAIN.
 RX MEDLINE=96311878; PubMed=8590002;
 RA Gosser Y.Q., Zheng J., Overduin M., Mayer B.J., Cowburn D.;
 RT "The solution structure of Abl SH3, and its relationship to SH2 in
 RT the SH(32) construct";

RL Structure 3:1075-1086(1995).
 RN [10]
 RP 3D-STRUCTURE MODELING OF SH3 DOMAIN.
 RX MEDLINE=95199229; PubMed=7892170;
 RA Pisabarro M.T., Ortiz A.R., Serrano L., Wade R.C.;
 RT "Homology modeling of the Abl-SH3 domain";
 RL Proteins 20:203-215(1994).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 57-218.
 RX MEDLINE=96398698; PubMed=8805595;
 RA Nam H.-J., Hasegawa W.G., Roberts T.M., Frederick C.A.;
 RT "Intramolecular interactions of the regulatory domains of the Bcr-Abl
 RT kinase reveal a novel control mechanism";
 RL Structure 4:1105-1114(1996).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS) OF 64-121.
 RX MEDLINE=96365516; PubMed=9698566;
 RA Pisabarro M.T., Serrano L., Wilmanns M.;
 RT "Crystal structure of the Abl-SH3 domain complexed with a designed
 RT high-affinity peptide ligand: implications for SH3-ligand
 RT interactions";
 RL J. Mol. Biol. 281:513-521(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; IA (SHOWN HERE) AND IB; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -1- DISEASE: PARTICIPATES IN A T(9;22)(3q41;q11) CHROMOSOMAL
 CC TRANSLOCATION THAT PRODUCES A BCR-ABL ONCOGENE RESPONSIBLE FOR
 CC CHRONIC MYELOID LEUKEMIA (CML), ACUTE MYELOID LEUKEMIA (AML), AND
 CC ACUTE LYMPHOBLASTIC LEUKEMIA (ALL).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/chronocancer/genes/ABL.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL, X16416; CAA34438.1; -;
 DR EMBL, M14752; AAB51561.1; -;
 DR EMBL, U07563; AAB60394.1; -;
 DR EMBL, U07563; AAB60393.1; -;
 DR EMBL, U07561; AAB60393.1; JOINED.
 DR EMBL, S69223; AAD14034.1; -;
 DR PIR, A25562; TVHA.
 DR PDB, 1AB2; 31-JAN-94.
 DR PDB, 1ABL; 01-NOV-94.
 DR PDB, 2ABL; 04-SEP-97.
 DR PDB, 1AWO; 28-JAN-98.
 DR PDB, 1BBZ; 25-NOV-98.
 DR Genem; HGNC:76; ABL1.
 DR MIM; 189980; -;
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR001455; Tyr_pkinase.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00109; TYRKINASE.
 DR Prodom; PD000001; Euk_pkinase; 1.
 DR Prodom; PD000066; SH3; 1.
 DR Prodom; PD000093; SH2; 1.


```

OY 844 AAGGCTCATGAGG-----TCGACGAGTGTGCTCCCTGGG 882
DB 617 gseraserpheargglumetaspglpProaspargly----- 632
OY 883 ATGAAGTGAGCTGCACCTCCAGATCCCTGTGGACACGACGAGACG 942
DB 633 -----AlasergluaspsaserArglulencysasnnglyProproAlaLeuth 649
OY 943 AAAATCATCATACCTCAACCTCAAGGAGTGGCCCTTAACACACCCACAGGCTTG 1002
DB 649 rseraspalaalagluProthLysSer----- 658
OY 1003 GATACCTCAGCTGTGTCACCTCTTGTGGCCGTGCTGATTAATAAGATTCCTAC 1062
DB 659 ----ProlysAlaserasnnglyAlaglyValProasnnglyAlaPheargglProglYas 677
OY 1063 CTGGCTTACGGGGCTGCGCATGGCGCTTGCGCTGTGTTCCGCTGGCGGGGACAC 1122
DB 677 nserglYpheargSerProHlsmet-----TpyLysLysSerSe 690
OY 1123 CCAAGACAGAGCTGCTCTACCTGTCTACACAGCAACAGGTCACAGTTCAGCATC 1182
DB 690 rthreuthrgrglseralaglyleualalalagluugluserglYmetSerSerly 710
OY 1183 GCGGATG-----AAGACGACGCGACAGACCCCTACCCAGTGC 1218
DB 710 sarphelauraserCysSerAlaserCysmetProHlsglyAlaArgspThrglu-- 729
OY 1219 AAGCCATGAGAGCTGTGACACAGCGGCTGTGCTGTGTACAGCAATGGCGGGCCCTC 1278
DB 730 -----TrrpArg--ServaThleuproArgspLeuproSerAlaglyLysglNph 746
OY 1279 CTTCATCATGAGCTGTGCTCCCTGAGATCTGACAGCGGCTGACCCCTACATGAGCC-- 1336
DB 746 easpseraserThrphneglyglYhlisLysSerclu-----LysProAlaLeupProAr 763
OY 1337 -----CC 1338
DB 763 glysArgThrsergluserArgserglugluValAlaLysSerThralaMetProLeupr 783
OY 1339 TCCATGCTTACCTGCTGTGTCAGCTGTGAGGACAGGGGAGAGAGTCTCTGGTGC 1398
DB 783 oolYtrpleuLysLysasn-----glugluAlaAlagluLuglyLph 797
OY 1399 CTGATGTGACAGGACCACTCTGTGTGATGTACACACTCCACACCTACAGCTGTGCTGC 1458
DB 797 elysAsphrgrgluserSerPro-glySerSerPro-----ProserLeuthrP 813
OY 1459 CGGACTCTCTGCGG-----GTCGCCAGCCCTCAGGAGGACATGTTTCCC 1503
DB 813 rolyLeuLeuArglArgluValThralaserProserSerLysLeuSerHlsgluG 833
OY 1504 GTGCG-----GCCCTTGACAGAGAACCCCGGACGACGACGACGACGAC 1551
DB 833 luAlaThrlYsglySerAlaserglYmetglYth- ProAlaThrlaAlaProAlaPro 852
OY 1552 CCAAGAGTCCCTGAGGAGGAGCTCATCCGCGAGCTGAGCATGTACAGTACAGAGCTG 1611
DB 853 ProserAsnlyValglYleuSerLysAlaser-----Serlndlunet 867
OY 1612 GGCACGACATCTGATCCACGAGAAATCACTGACTGACTGCTCCATGCTCTAC 1671
DB 868 -----ArgvalArgArgHlslshlsser----- 875
OY 1672 TCCCTCATCCCAACCCCGCAGGCTGCACG----- 1701
DB 876 SergluserProglYArgspLysglYargluAlaAlaLysLeuLysProAlaProProPro 895
OY 1702 -----TCCCTCATCCAGCTCCCGACGCGCCGACGACCACT 1734
DB 895 ProProAlaCysThrlYsglyAlaLysLysProAlaLysLysProserSerlndlunet 915

```

```

OY 1735 TCTCCAGTGTGCTCTTCTCCACGAGCTGCAGAGAC-----TCAGACATGCTACATACG 1788
DB 916 GluAlaLysglYProthLysCysThrSerLeuAlaMetAlaValAlaValasnThr 935
OY 1789 -----CCCGGTGCTGCTCCGACAGGCTGTGACATGACCTGACCCCATGAC 1836
DB 936 AspProthLysAlaLysPro-ProglYgluglyLeuArgLysProValProProSerVa 955
OY 1837 GGGGACCTTCAGCAGCAGGCGGTGAGATGTATGCTCATGTGCGCTGAGAGATTCGAAACC 1896
DB 955 lProLysProGln-----SerThralaLysProProglYthrProthSerProValSer-- 973
OY 1897 TGGTCCCAAGCGCGGTGAGATGTATGCTCATGTGCGCTGAGAGATTCGAAACC 1956
DB 974 ----ThrProserThralaProAlaProSerProLeuAlaLysglYasnGln-----GlnPr 990
OY 1957 CAGCGGCGCGGACGTCATGCGG 1978
DB 990 oSerSerAlaAlaPhellePro 997

```

RESULT 6

3K9_HUMAN STANDARD; PRT; 394 AA.

AC P80192; 1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase 9 (EC 2.7.1.-) (Mixed
 DE kinase kinase 1) (Fragment).
 GN MAP3K9 OR MLK1 OR PRKEL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid:9606;
 RN [1]

RC TISSUE=Colon epithelium;
 RA MEDLINE=93238756; PubMed=8477742;
 RX Dotzow D.S., Devereux L., Dietzsch E., de Kretser T.;
 RT "Identification of a new family of human epithelial protein kinases
 RT containing two leucine/isoleucine-zipper domains.";
 RL Eur. J. Biochem. 213:701-710(1993).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF
 CC COLORECTAL, BREAST AND OESOPHAGEAL ORIGIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC MAP KINASE KINASE KINASE SUBFAMILY.

DR PIR: S32467; S32467.
 DR PIR: J00229; J00229.

DR HSSP: P12931; IFMK.
 DR GeneW: HGNC:6861; MAP3K9.

DR MIM: 600136;
 DR InterPro: IPR000719; Euk_Pkinase.

DR InterPro: IPR002290; Ser_Thr_Pkinase.
 DR InterPro: IPR001245; Tyr_Pkinase.

DR Pfam: PF00068; Pkinase; 1.
 DR ProDom: PD000001; Euk_Pkinase; 1.

DR SMART: SM00219; TYKRC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS00106; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
 KW ATP-binding.

FT NON_TER 1 1
 FT DOMAIN 3 271 PROTEIN KINASE.
 FT NP_BIND 9 17 ATP (BY SIMILARITY).

FT BINDING 30 30 ATP (BY SIMILARITY).
 FT ACT_SITE 127 127 BY SIMILARITY.

FT DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).
 FT DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).
 FT DOMAIN 354 368 ARG/LYS-RICH (BASIC).

SU SEQUENCE 394 AA; 44975 MW; DBE10R7D319477PDB CAC64;

Alignment Scores:


```

OY 1075 GGCCCTGCGAGGAGGCTGTGGCTGTGTTTCCCGTGGTGGGAGCACCACCAAGGACAGC 1134
DB 554 ALeuSerGlyValGlyLeu----- 560
OY 1135 TGCCTCTACCTGTGCTGCACACACACAGTCCAGATTCAGATCGCGATGAGAC 1194
DB 561 ----ProGlyCys-----ProlGlySerProProSerProGlyArgSerArg-- 574
OY 1195 GCACGGCAGAACCCCTACCCAGTGAAGCCATGAGGTGGTCAACAGCGGCTCTAGATC 1254
DB 575 -ArgGlySerThrArgHisArg-----LysAlaSerAlaIly 586
OY 1255 TGGTACAGAAAGGCGCGGCGCTCTGTGCATCGACTGCTCCCTGGAGATCGACAG 1314
DB 586 sGlySerCysGlyAspLeuProGlyLeuArgAlaAlaLeuProHisGluProGlyG1 606
OY 1315 CGGCTGAGAGCCCT-----ACATGGCCC-----CTTCATG 1344
DB 606 YLeuGlySerProGlyGlyLeuGlyValGlyProSerAlaTrpAlaCysProProAl 626
OY 1345 GTTACGT-----CA 1353
DB 626 ALeuArgGlyLeuHisAspLeuLeuArgIlyMetSerSerSerProAspLe 646
OY 1354 GTCGTGTGAGCTGTGAGGAGGAGGAGAGAGTGTGTGCTGTGCTGGATGACA----- 1408
DB 646 uLeuSerAlaAlaLeuGlyAlaArgGlyArgGlyAlaThrGlyGlyAlaArgAspProG1 666
OY 1409 -AGGCCAACCTCTGTGTGTGTATGATACATCCACCCATCCACCTGTGTGCCGCTACTTC 1467
DB 666 YSerProProProGlnGlyAspThrProProSerGlnGlySerAlaProGlySerTh 686
OY 1468 TGGCG-----GGTCCCGCAGCGCCCTCAGGAGACATGTTTCC 1503
DB 686 rSerProAspSerProGlyGlyAlaIlyGlyLupProPro-----Pr 701
OY 1504 GTGCGAGCCCTTGACACAGGAMCCCCGAGCAGCACACAGCGCCACCAAGGTGCT 1563
DB 701 oValGlyPro-GlyGlyGlyVal----- 708
OY 1564 GAGGGGAGCTCATGCGGAGCTGAGCATCATGATACATGAGAGAGTGGGACCGCATC 1623
DB 709 --GlyLeuLeuGlyThrGlyArgGlyGlyThrAlaGlyArgGlyIlyAlaArgAla 726
OY 1624 CTGATCCACACAGAAATCACTCACTGACTGCTCCATGCTCCATGCTCCATGCTCCCA 1683
DB 727 -----GlySerGlnHisLeuThrProAlaAlaLeuLeu-TyrArgAlaAlaVal 742
OY 1684 CCCCGCAGGCTGCGAGGTCCTCCAGCCCTCCAGCTCCCGACAGCAAGTCTTCCACT 1743
DB 743 ThrArgSerGlnLysArgGlyIlyLeuSerSerGlnGlnGlnGlyValAlaAspSerGlu 762
OY 1744 GTGCTTTTCCACGAGCTGCGAGGACTGACACATGCTACATACGCCCGGCTGCTGCCCTC 1803
DB 763 ValGluLeu-ProProSerGlnArg----- 770
OY 1804 GACAGGTGAGACATGACTGACCCCATGAGCGGGAGACCTTCAGCCAGACCTGAG 1863
DB 771 -----TrpProGlnGlyProAspMet----- 777
OY 1864 GCCGTGAGATCTGCGCCCTCAGAGACCTATTGGGTCCCGCAGCGCGGTGAGATGTT 1923
DB 778 -----ArgLysSerLeuSerThrPheSerSerGluAsnProSerAspValGlu----- 793
OY 1924 ATGCTGATTGGCTTGAGAGAGATTGTGAAGCCAGCGGGCGGAGATCTTCCGCTTGA 1983
DB 794 -----GluGlyThrAlaSerGluProSerProSerG1 804
OY 1984 AAGCCCGAG 1993
DB 804 YThrProGlu 807

```

```

RESULT 5
ABL_MOUSE STANDARD; PRT: 1123 AA.
ID ABL_MOUSE AC P00520; Q61252; Q61253; Q61254; Q61255; Q61256; Q61257; Q61258;
AC Q61259; Q61260; Q61261; P97896;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150)
DE (c-ABL).
GN ABL1 OR ABL.
OS Mus. musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RX MEDLINE=88068561; PubMed=3317402;
RA Oppl C., Shore S.K., Reddy E.P.;
RT "Nucleotide sequence of testis-derived c-abl cDNAs: implications for
RL testis-specific transcription and abl oncogene activation.";
RN [2]
RP SEQUENCE OF 1-187 FROM N.A. (ISOFORMS I TO IV).
RX MEDLINE=95394474; PubMed=7665185;
RA Chisoe S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burtan D.,
RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jian L., Ma Y.,
RA McLaury H.-J., Pan H.-O., Sarhan O.H., Toth S., Wang Z., Zhang G.,
RA Heisterkamp N., Groffen J., Roe B.A.;
RT "Sequence and analysis of the human ABL gene, the BCR gene, and
RN regions involved in the Philadelphia chromosome translocation.";
RL Genomics 27:67-82(1995).
RN [3]
RP SEQUENCE OF 85-182 FROM N.A.
RX MEDLINE=84106840; PubMed=6319018;
RA Wang J.Y.-J., Ledley F., Goff S., Lee R., Groner Y., Baltimore D.;
RT "The mouse c-abl locus: molecular cloning and characterization.";
RL Cell 36:349-356(1984).
RN [4]
RP ALTERNATIVE SPLICING.
RX MEDLINE=88202920; PubMed=3283651;
RA Beardards A., Paskind M., Baltimore D.;
RT "Four murine c-abl mRNAs arise by usage of two transcriptional
RN promoters and alternative splicing.";
RL Oncogene 2:297-304(1988).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 61-121.
RX MEDLINE=95393198; PubMed=7664083;
RA Musacchio A., Saraste M., Williams M.;
RT "High-resolution crystal structures of tyrosine kinase SH3 domains
RN complexed with proline-rich peptides.";
RL Nat. Struct. Biol. 1:546-551(1994).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; THE MYRISTOYLATED C-ABL PROTEIN
CC WAS REPORTED TO BE NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; I (SHOWN HERE), II, III AND
CC IV, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).

```

CC -1- COFACTOR: Magnesium.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -1- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
 CC testis, gastrointestinal tract, stomach, liver and pancreas.
 CC within the nervous system, predominantly expressed in neurons and
 CC enriched in synaptic terminals.
 CC -1- PFM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC
 CC EMBL; U14636; AAA57280.1; -
 CC DR HSSP; P12931; IFMK
 CC DR MGD; MGI1346881; Map3k12.
 CC DR InterPro; IPR000719; Euk_Pkinase.
 CC DR InterPro; IPR004040; Ser_Pkinase.
 CC DR Pfam; PF00069; Pkinase; 1.
 CC DR ProDom; PD000001; Euk_Pkinase; 1.
 CC DR SMART; SM00221; STYK; 1.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
 CC KW Phosphorylation; Magnesium; Membrane.
 CC FT DOMAIN 158 399 PROTEIN KINASE
 CC FT NP_BIND 164 172 ATP (BY SIMILARITY).
 CC FT BINDING 185 185 ATP.
 CC FT ACT_SITE 269 269 BY SIMILARITY.
 CC FT DOMAIN 56 62 POLY-GLY.
 CC FT DOMAIN 668 671 POLY-PRO.
 CC FT DOMAIN 698 701 POLY-PRO.
 CC FT DOMAIN 753 758 POLY-GLU.
 CC FT MUTAGEN 185 185 K->A; NO CATALYTIC ACTIVITY.
 CC FT MUTAGEN 192 192 E->A; NO CHANGE.
 CC FT CONFLICT 18 18 V -> A (IN REF. 2).
 CC FT CONFLICT 28 29 KL -> NY (IN REF. 2).
 CC FT CONFLICT 382 382 S -> T (IN REF. 2).
 CC FT CONFLICT 494 495 EQ -> DE (IN REF. 2).
 CC FT CONFLICT 517 517 N -> D (IN REF. 2).
 CC FT CONFLICT 794 794 E -> G (IN REF. 2).
 CC SQ SEQUENCE 888 AA; 96083 MW; CFCFLD34AF89ABB CRC64;
 Alignment Scores:
 Pred. No.: 3.23e-08 Length: 888
 Score: 290.50 Matches: 177
 Percent Similarity: 33.68% Conservative: 81
 Best Local Similarity: 23.11% Mismatches: 259
 Query Match: 7.16% Indels: 249
 DB: 1 Gaps: 32
 US-09-836-392-8_COPY_22_2205 (1-2184) x M3KC_MOUSE (1-888)
 QY 28 GCCATGAACTTCGCGAGTTCGGGAG---GAGCGACAGATGCTGACGGGCTGCAG 84
 DB 183 AAlaValLsLysValAlrGAspleuLysgluThrAspLleLysHisLeuArGlySLeuLys 202
 QY 85 CACCCCTCATGCTGCGCTCATCGGCATCAGCATCCACCCGCTGCTCGCTCGCC----- 138
 DB 203 HlSProAsnLleLleThrPheLysGlyValLysThrGlnAlaProCysTyrCysIleLeu 222
 QY 139 CTGGAGCTGGCGCGCTGACGACCTTCACACACCGCTGCTGCTGCGAGAACGCGACAGATTCT 198
 DB 139 CTGGAGCTGGCGCGCTGACGACCTTCACACACCGCTGCTGCTGCGAGAACGCGACAGATTCT 198

DB 223 MetGluPheCysAlaGlnGlyLeuLeuTyrGluValLeuAlaGlyArg----- 239
 QY 199 TCCATTATACCCCTGGAGACATGCTCACCCCAAAATAGCTTACCAATGCTCGGC 258
 DB 240 -----ProValThrProSerLeuLeuValAspTrpSerMetGlyLleAlaGlyGly 256
 QY 259 CTGGCCTTACCTGGCACACAGAAAACATCATCTTCTGTGACSTGGAAGTGGACACATTTCT 318
 DB 257 MetAsnTyrLeuHisLysLleLysLleLleHisArgArgLysSerProAsnMetLeu 276
 QY 319 GTGTGCTCCCTGACGTGACAGAGACCATCAACATCAAGTATGATGATGACGATTGCG 378
 DB 277 HleThrTyrAspAspVal-----ValLysLleSerAspPheGlyThrSer 291
 QY 379 AGCGAG---TCATTTCATAGAGGCGCCCTAGCGCTGAGGCGACSTCTGCTACAGGCG 435
 DB 292 LysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThrValAlaIlePheAla 311
 QY 436 CCAGAGATCAGGCGCTCGCATTTGATATGATGAGAAGATGATATGCTCTCATGAGATG 495
 DB 312 ProGluValLleArgAsnGluProValSerGluLysValAspLleThrSerPheGlyVal 331
 QY 496 GTGCTCTACAGATTGCTGTCAGAGACAGGCGCT----- 528
 DB 332 ValLeuThrPgluLeuLeuThrLysLleLysLleProTyrLysAspValAspSerLleAla 351
 QY 529 -----GCAGTGGCGCACCCAGCCAGCTCCAGATTGCCAAGAAAGCTGTCCAGGCGATCCGC 582
 DB 352 HleThrPgluValGlySerAsnSerLeuHisLys----- 362
 QY 583 CCGGTTCTGGGGCAGCCGAGAGATGCAATTCGCGCGATCAGCGCTCATGATGAG 642
 DB 363 ProValProSerSerCysProAsp-----GlyPheLysLleLeuLeuArgLys 378
 QY 643 TGCTGGGACATCAAGCCAGAGAACGAGCGG-----CTGGCCCTGTCG 684
 DB 379 CysThrPasnSerLysProArgAsnArgProSerPheArgLysLleLeuLeuHisLysAsp 398
 QY 685 GTGTGAGCGACAGATCAAGAGCCGACTTTGGCCACTTCATGATGAACTGCTGTGGG 744
 DB 399 HleAlaSerAla-----AspValLeuSerThr 407
 QY 745 AAGCAGACAGCTTCTTTCATCCAGGGCGAG-----GAGTAC 783
 DB 408 ProGlnGluThrTyrPheLysSerGlnAlaGluTyrArgLysLysLysHisPhe 427
 QY 784 ACCGTGCTTTGGGATGGA----- 804
 DB 428 GluLysLleLysSerGluLysLysThrCysLeuHisArgLeuGluGluLeuValMetArg 447
 QY 805 -----AAGAGAGATCCAGAGAC----- 822
 DB 448 ArgArgGluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 467
 QY 823 -----TACACGCTGTGAACACA-----GAG 843
 DB 468 ArgAlaAsnAsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 487
 QY 844 AAGGCGCTCATGAGAGTGCAGAGATGTGC-----TGCCCTGGGATGAGAGTGCAGC 894
 DB 488 ArgGluLeuLeuArgArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys 507
 QY 895 TGCCAGCTCCAGATCCAGATCCCTGTGAGACAGCCAGCCAGGACAGACAAATCTATCATC 954
 DB 508 -HlSProSerArgGlyLeuLeuHisGlyAsnThrMetGluLysLysLysLysLysLysLys 527
 QY 955 TACACCTCAAGAGGATGTGCCCTTAAACACACCCCAAGCGCTGTGATCTCCAGCT 1014
 DB 527 nValProGlnLysLeuSer-----ProHisSerLysArgProAspLleLeuLysThr 544
 QY 1015 GTGCTACACTGCTTCTTGCCCGCTGCTGTTATTAAGAAATCTCTACTGCTGTAGCG 1074
 DB 544 rGluSer-----LeuLeuProLysLeuAspAlaL 554


```

Db 408 PROGLINLTHRTGPRHELYSSERGLNLAGLTPRAGSLUGLULVALYLSLEUNHSPE 427
QY 784 ACCGTGGTGTGGATGGA----- 804
Db 428 GLULYLSLEUNHSERGLUGLULYTHCYLSLEUNHSARGLEUGLUGLULVALMETARG 447
QY 805 ---AAGAGAGAGGAGGAGAGC----- 822
Db 448 ARGARGGLUGLULYLSLEUNHSARGLEUNHSARGLEUNHSARGLEUNHSARGLEUNHS 467
QY 823 -----TACAGGTGGTGAACACA----- 843
Db 468 ARGALASNAHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 487
QY 844 AAGGGCTGTAGAGGTGAGAGAGATGTC-----TGCCTGGATGAAGGTGAGC 894
Db 488 ARGGLULYLSLEUNHSARGLEUNHSARGLEUNHSARGLEUNHSARGLEUNHSARGLEUNHS 507
QY 895 TGCAGCTCCAGGTCCAGAGATCCCTGTGGACAGCCAGGAGGAGAGAGAGAGAGAGATC 954
Db 508 -HSTHNSERARGSERLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 527
QY 955 TACAGCTCCAGGAGGAGATGTCCTTAACACACACCCAGGAGGAGGAGGAGGAGGAGT 1014
Db 527 NVALPROGLINLTHLEUNHSER-----PROHISSELYSARGPROSPRLEUNHSLEUNHS 544
QY 1015 GTCGTACCTGCTTGTGGCGCTGCTGTATTAAAAAGATTCCTACCTGGCTTACGC 1074
Db 544 RGLUSER-----LEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 554
QY 1075 GGCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGC 1134
Db 554 ALEUSERGLYVALGLYLEU----- 560
QY 1135 TGCCTACCTGCTGCTACACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1194
Db 561 ---PROGLINLTHLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 578
QY 1195 GCAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1254
Db 578 RAGHNSHLYG-----LYSLASERLALY 586
QY 1255 TGGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1314
Db 586 SGLYSERGLYASPRLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 606
QY 1315 CGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1344
Db 606 YLEUGLYSERPROGLYGLYLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 626
QY 1345 GTTAGCT-----CA 1353
Db 626 ALEUAGGLYLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 646
QY 1354 GTCGTGTCAGCTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1408
Db 646 ULEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 666
QY 1409 -AGGCAACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1467
Db 666 YSERPROSPRLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 686
QY 1468 TGGC-----GGGTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1503
Db 686 TSEPRPROASPRSERPROGLYGLYVALYLSGLYGLYUPTORPORO-----P 701
QY 1504 GTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1563
Db 701 OVALGLYPTRO-GLYGLYGLYVAL----- 708
QY 1564 GAGGGGAGTCCATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1623

```

```

Db 709 --GLYLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 726
QY 1624 CTGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1683
Db 727 -----GLYTYRGLINHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 742
QY 1684 CCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1743
Db 743 THRARGSERGLINHSARGGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 762
QY 1744 GTCCTTTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1803
Db 763 VALGLULEN-PROSPRLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 770
QY 1804 GACAGGTCTGAGCATACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1863
Db 771 -----TRPPROGLINLTHLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 777
QY 1864 GCGGTGAGATCCGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1923
Db 778 -----ARGINLSERLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 793
QY 1924 ATCTGATTCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1983
Db 794 -----GLYGLYTHRALASERGLUPROSPRLEUNHSLEUNHSLEUNHSLEUNHSLEUNHS 804
QY 1984 AAGCGCGAG 1993
Db 804 YTHPRGLIN 807

RESULT 4
M3KC.MOUSE STANDARD: PRT: 888 AA.
AC 060700: P70286;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
DE (leucine-zipper protein kinase) (ZPK) (dual leucine zipper bearing
DE kinase) (DLK)
GN MAPK12 OR ZPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Brain;
RX MEDLINE=95074107; PubMed=7983011;
RA Holman L.B., Merritt S.E., Fan G.;
RT Identification, molecular cloning, and characterization of dual
RT leucine zipper bearing kinase. A novel serine/threonine protein kinase
RT that defines a second subfamily of mixed lineage kinases."
RT J. Biol. Chem. 269:30808-30817(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
RX MEDLINE=96365388; PubMed=8769555;
RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Groulx G.;
RT "Cell-specific expression of the ZPK gene in adult mouse tissues."
RT DNA Cell Biol. 15:631-642(1996).
RN [3]
RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
RX MEDLINE=96279269; PubMed=8663324;
RA Mata M., Merritt S.E., Fan G., Yu G.G., Holman L.B.;
RT Characterization of dual leucine zipper-bearing kinase, a mixed
RT lineage kinase present in synaptic terminals whose phosphorylation
RT state is regulated by membrane depolarization via calcineurin."
RT J. Biol. Chem. 271:16888-16896(1996).
CC -I- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -I- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.

```

Db 728 SerGluValGluLeuThrSerSerGlnArgTyrProGlnSerLeuAsnMetArgGln--- 746
 Qy 1639 TCACCTACGACTGACTGCTCCATGCTCCCTCCATCCCTCCACCCGAGGCTGCC 1698
 Db 747 -----SerLeuSerThrPheSerSerGlnAsnProSerAspGlyGlu 760
 Qy 1699 AGCTCCCTCAGCCCTCCAGCTCCCTCCAGCAAGCTTCCTCAGTGTCTTCACCC 1758
 Db 761 GluGlyThrAlaSerGluPro---SerProSerGlyThrProGluValGly---SerThr 778
 Qy 1759 GACCTGGAGGAGCTGACATGCTACATACGCGGCTGCTCCGACAGCTGAGCT 1818
 Db 779 AsnThrAspGluArgProAspGluArgSerAspAspMetCysSerGlnGlySerGluIle 798
 Qy 1819 GACCTGACCCCTGAGCGGAGG 1842
 Db 799 ProLeuAspProProProSerGlu 806 --
 RESULT 3
 M3KC_RAT STANDARD: PRT; 888 AA.
 AC 063796;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
 DE (MAPK-upstream kinase) (MUK).
 GN MAPK12 OR MUK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96226099; PubMed=8637721;
 RA Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.;
 RT "Activation of the JNK pathway by distantly related protein kinases,
 MEK and MUK";
 RL Oncogene 12:641-650(1996).
 CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein - ADP + a phosphoprotein.
 CC -1- COFACTOR: Magnesium.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
 CC similarity).
 CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: D49785; BAA08621.1; -
 DR HSSP: P12931; 1FMK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00221; STYK; 1.
 DR PROSITE: PS00107; PROTEIN KINASE AMP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Magnesium; Membrane.

FT DOMAIN 158 399 PROTEIN KINASE:
 FT NP_BIND 164 172 ATP (BY SIMILARITY).
 FT BINDING 185 185 ATP (BY SIMILARITY).
 FT ACT_SITE 269 269 BY SIMILARITY.
 FT DOMAIN 56 62 POLY-GLY.
 FT DOMAIN 668 671 POLY-PRO.
 FT DOMAIN 698 701 POLY-PRO.
 FT DOMAIN 753 758 POLY-GLU.
 SQ SEQUENCE 888 AA; 96307 MW; 52AD964006BAE149 CRC64;
 Alignment Scores:
 Pred. No.: 2, 01e-08 Length: 888
 Score: 294.50 Matches: 177
 Percent Similarity: 33.55% Conservative: 80
 Best Local Similarity: 23.11% Mismatches: 260
 Query Match: 7.26% Indels: 249
 Gaps: 31
 DB: 1
 US-09-836-392-8_COPY_22_2205 (1-2184) x M3KC_RAT (1-888)
 Qy 28 GCCATGAGAACTCTCCAGTTCGCGAG---GAGGCCAGCATGCTGCACCGCTGCAG 84
 Db 183 AlaValLysLysValAlaArgAspLeuLysGluThrAspIleLysHisLeuArgLysLeuLys 202
 Qy 85 CACCCCTGATGCTGCGCTCATCGGATCAGCATCCACCCGCTGCTGCC----- 138
 Db 203 HisProAsnIleIleThrPheLysGlyValCysThrGlnAlaProCysTyrCysIleLeu 222
 Qy 139 CTGGAGCTGCGCGCTGACAGCTCAACACCGCTGCTCCGAGAACCCAGAACTTCT 198
 Db 223 MetGluPheCysAlaGlnGlyLysLeuTyrGluValLeuAlaGlyArg----- 239
 Qy 199 TCCTTATACCTCGGAGACATGCTCACCCCAAAATAGCTTACCATCGATCGCTCGGC 258
 Db 240 -----ProValThrProSerLeuLeuValAspTyrSerMetGlyIleAlaGlyGly 256
 Qy 259 CTGGCTTACCTGCACAGAAATATCATCTTCTGACCTGAGTGCAGACATCTTG 318
 Db 257 MetAsnTyrLeuHisLeuHisLysIleIleHisArgAspLeuLysSerProAsnMetLeu 276
 Qy 319 GTGTGCTCCCTGACCTCAAGAGACATCAACATCAACCTATCTGACATCGGATTGG 378
 Db 277 IleThrTyrAspAspVal-----ValLysIleSerAspPheGlyThrSer 291
 Qy 379 AGGCAG---TCATTTCATGATGAGCGCCCTAGCGCTGAGCGCATCTGCTACAGGCC 435
 Db 292 LysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThrValAlaTyrMetAla 311
 Qy 436 CCAGAGATCAGGCTCGCATTTGATATGATGAGAGAGATGATATGCTCATGGAATG 495
 Db 312 ProGluValIleArgAsnGluProValSerGluLysValAspIleTyrSerPheGlyVal 331
 Qy 496 GTGCTTACGATGCTGCTGCAGACAGCCCT----- 528
 Db 332 ValLeuThrGluLeuLeuThrGlyGlyIleProTyrLysAspValAspSerSerAlaIle 351
 Qy 529 -----GCACTGGGCGACACACAGCTCCAGATTGCCAAGAGTGTCCAAAGGCGATCCG 582
 Db 352 IleTyrGlyValGlySerAsnSerLeuHisLeu----- 362
 Qy 583 CCGGTTCTGGGCGACCGGAGAGATGCAATTCGCGCATCGAGCGCTCATGATGAG 642
 Db 363 ProValProSerSerCysProAsp-----GlyPheLysIleLeuLeuArgGln 378
 Qy 643 TCGTGGACACTAAGCCAGAGAACGACCG-----CTGGCCCTGTGG 684
 Db 379 CysTyrAsnArgLysProArgTargAsnArgProPheArgGlnIleLeuLeuHisLeuAsp 398
 Qy 685 GTGGTATGACCAATGAAGACCGGACTTTTGCACCTTATGATGATGATGATGCTGG 744
 Db 399 IleAlaSerAla-----AspValLeuSerThr 407
 Qy 745 AAGCAGACAGCTTCTTCTCATCCAGGCCAG-----GAGTAC 783

SQ SEQUENCE 859 AA: 93188 MW: 05209792C5C6F05 CRC64:
 Alignment Scores:
 Pred. No.: 1.78e-08 Length: 859
 Score: 295.50 Matches: 168
 Percent Similarity: 34.27% Conservative: 89
 Best Local Similarity: 22.40% Mismatches: 257
 Query Match: 7.28% Indels: 236
 DB: 1 Gaps: 31
 US-09-836-392-8_COPY_22_2205 (1-2184) x M3KC_HUMAN (1-859)
 QY 28 GCCATGAAACATCTCTCCAGATCCGGCAG---GAGCCAGATGCTGCACCGCTGCAG 84
 DB 150 AAlaValLysLysValAlaArgLysLeuLysGluThrAspLLeuHisLeuArgLysLeu 169
 QY 85 CACCCCTGATCGGCGCTCATCGGATCAGATCCACCCCTGCTGCTGCC----- 138
 DB 170 HisProAsnLLeuLeuThrPheLysGlyValCysThrGlnAlaProCysLysLLeu 189
 QY 139 CTGAGACTCGCCGCTGACAGCCTCAACACCGCTGCTGCAGAACGCCAGATCT 198
 DB 190 MetGluPheCysAlaGlnLysLeuLysGluValLeuAlaArgLysArg----- 206
 QY 199 TCCTTTAAACCCCTGGAGACATGCTCAACCAAAATAGCTTACAGATCCGCTGGCC 258
 DB 207 -----ProValThrProSerLeuValAlaThrPseMetGlyLLeuAlaGly 223
 QY 259 CTGGCTTACCTCCACAGAAACATCATCTCTGTGACCTGAAGTGGACAACTTCG 318
 DB 224 MetAsnLysLeuHisLeuHisLysLLeuHisArgLysLeuLysSerProAsnMetLeu 243
 QY 319 GTGTGCTCCCTTACGCTCAAGAGACATCAATCAATCAAGCTATGACTAGCGGATTCG 378
 DB 244 LLeuThrLysAspAspAl-----ValLysLLeuSerAspPheLysThrSer 258
 QY 379 AGCCAG---TCATTCCATGAGGCGCCCTAGCGCTGGAGGCACTCTGGTACAGGCC 435
 DB 259 LysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThrValAlaThrPheAla 278
 QY 436 CCAGAGATCAGGCTCGCATGTATATATGATGAGAAGGTAGATGTCCTCATGGAAATG 495
 DB 279 ProGluValLLeuArgAsnGluProValSerGluLysValAspLLeuThrPseMetGlyVal 298
 QY 496 GTGCTCTACGATCTCTGTACAGACAGCGCCT----- 528
 DB 299 ValLeuThrGluLeuLeuThrGlyGluLLeuProLysLysAspValAspSerSerAlaLeu 318
 QY 529 -----GCACCTGGCCACACCAAGCTCCAGATTCACAGAGCTGTCCAGGGCATCCGC 582
 DB 319 LLeuThrGlyValGlySerAsnSerLeuHisLeu----- 329
 QY 583 CCGGTTCTGGGACCGGACGAGATGACGTCCGGCGACTCAGCGGCTCATGATGAG 642
 DB 330 ProValProSerSerCysProAsp-----GlyPheLysLLeuLeuArgGln 345
 QY 643 TCGTGGGACACTAACCCACAGAGAGCGACCG-----CTGGCCCTGTGC 684
 DB 346 CysTrpAsnSerLysProArgAsnArgProSerPheArgGlnLLeuLeuHisLeuAsp 365
 QY 685 GTGGGAGCCAGATGAAGAGACCCGACTTTGCGACCTTATGATGAACTGTGCTGTGG 744
 DB 366 LLeuAlaSerAla-----AspValLeuSerThr 374
 QY 745 AAGCAGACAGCTTCTTCATCCAGGGCGAG-----GAGTAC 793
 DB 375 ProGlnGluThrLysPheLysSerGlnAlaGluThrArgGluValLysLeuHisPhe 394
 QY 784 ACCGCTGCTTTGGGATGA----- 804
 DB 395 GluLysLLeuLysSerGluLysLysCysLeuHisArgLeuGluGluLeuValMetArg 414
 QY 805 ---AAGAGAGAGTCCAGGAC----- 822

DB 415 ArgArgGluGluLeuArgHisAlaLeuAspLLeuArgGlnHisLysLysLysLys 434
 QY 823 -----TACACGGTGGTGAACACA-----GAG 843
 DB 435 ArgAlaAsnAsnLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 454
 QY 844 AAGGCTCATGAGAGGTGCAGAGATGTC-----TGCCTGGATGAGGTGAGC 894
 DB 455 ArgGluLeuLeuArgGluGluLysLysLysLysLysLysLysLysLysLysLys 474
 QY 895 TGCAGCTCCAGCTCCAGATCCCTGTGGACAGCCAGCCAGCCAGCCAGCCAGCCAGCC 954
 DB 474 HisProSerArgGlyLeuLeuHisGlyAsnThrMetGluLysLLeuLysLysLys 494
 QY 955 TACACCTCAAGGAGATGCGCCCTTAACACACCCCAACAGCCCTGGATCTCCAGCT 1014
 DB 494 nValProGlnAsnLeuSer-----ProHisSerGlnArgProAspLLeuLysAl 511
 QY 1015 GTGCTCACCTGCTTGGCCGCTGCTGCTTATTAAGAAATCTTACCTGCTTACGG 1074
 DB 511 agLuser-----LeuLeuProLysLysAspAlaAl 521
 QY 1075 GCGTCCCGGATGGCTTGTGCTGTGTTCCGCTGCTGCGGACCCCAAGAGACAGC 1134
 DB 521 AluSerGlyValGlyLeu----- 527
 QY 1135 TGTCTCTACCTGTGCTCACACACAGCCAGCCAGCTCAAGTTCAGATCGGATGAAGC 1194
 DB 538 ---ProGluLysProLysAlaProProSer---ProGluArgSerArgArgGlyLys 545
 QY 1195 GCACGCGAAGACCCCTTACCCAGTGAAGGCGATGAGGTGCTCAACAGCGGCTCTGAGGTC 1254
 DB 545 LArgHisArg-----LysAlaSerAlaLys 553
 QY 1255 TGTGACAGCAATGGCGCGGCTCTGTGCATGACGTGCTCCCTGGAGATCTGCAGG 1314
 DB 553 SerLysSerCysGluAspLeuProGluLeuArgThrAlaValProHisLysLysLys 573
 QY 1315 GCGCTGGAGCCCT-----ACATGGGCC-----CTGCCATG 1344
 DB 573 YProLysSerProGluGlyLeuGlyLysLysLysLysLysLysLysLysLysLys 593
 QY 1345 GTTACGT-----CA 1353
 DB 593 AluLeuArgLysLeuHisLysAspLeuLeuLeuArgLysMetSerSerSerProAspLe 613
 QY 1354 GTGCTGTGACGCTGTGAGGCGACAGGAGAGAGTGTGCTGTGCTGCTGCTGCTGCTG 1407
 DB 613 uLeuSerAlaAlaLeuGlySerArgGlyArgGlyAlaThrGlyGlyAlaGlyAspProG 633
 QY 1408 AAGCCCACTCTGTGATGTACACCTCCACCACTTACACCTGTGTCGCGGACTTC 1467
 DB 633 YSerProProProAlaArgLysAspThrProProSerGluLysSerAlaProGlySerTh 653
 QY 1468 TGGC-----GGTCCCAAGCCCTCAGGAGACATGTTCC 1503
 DB 653 rSerProAspSerProGlyAlaLysLysLysLysLysLysLysLysLysLysLys 668
 QY 1504 GTGCGCCCTTGGACAGGAGAC-----Pr 668
 DB 668 oValGlyPro-GlyGluGlyValGlyLeuLeuGlyThrGlyArgGluGluLysThrSerGly 688
 QY 1527 -----CCGGGACCCAGCCACACAGCC----- 1548
 DB 688 rGlyGlySerArgAlaLysSerGln-HisLeuThrProAlaAlaLeuLeuLysArgAla 707
 QY 1549 -----AACCAAGTGCCTGAGGAGGAGCTCATC 1578
 DB 708 AAlaValThrArgSerGlnLysArgGlyLysSerSerGluGluGluGluGluValAsp 727
 QY 1579 GCGAGCTGAGCATCATGTACAGTGAAGAGCTGGGACAGCAGATCTGATCCACAGAA 1638

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 15, 2003, 05:35:32 ; Search time 20 Seconds
(without alignments)
9058.430 Million cell updates/sec

Title: US-09-836-392-8_COPY_22_2205
Perfect score: 4058
Sequence: 1 atgcctgagcagcctgcggc.....ctgcactcgaagaagaag 2184

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 7.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 segs, 41476328 residues
Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2.1/USPRO/US9886332/runat_08042003_090327_22769/app.query.fasta.1.2375
-DB=SwissProt_40 -OPMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcr -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9886332 cgn2.1.16.ernat.08042003_090327_22769 -NCPU=6 -ICPU=3
-NO_XMAP -NO_XMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	312.5	7.7	954	M3KA_HUMAN	Q02779 homo sapien
2	295.5	7.3	859	M3KC_HUMAN	Q12852 homo sapien
3	294.5	7.3	888	M3KC_RAT	Q63796 rattus norv
4	290.5	7.2	888	M3KC_MOUSE	Q60700 mus musculu
5	280	6.9	1123	ABLI_MOUSE	P00520 mus musculu
6	279.5	6.9	394	M3K9_HUMAN	P80192 homo sapien
7	273.5	6.7	746	ABL_MLVAB	P00521 abelson mur
8	269	6.6	1130	ABLI_HUMAN	P00519 homo sapien
9	268	6.6	660	YH1_EBV	P03181 epstein-bar
10	265	6.5	660	YH1_EBV	P03181 epstein-bar
11	261	6.4	1050	ULK1_HUMAN	O75385 homo sapien
12	257	6.3	410	KYR2_DICDI	P18161 dictyostell
13	255.5	6.3	630	ULK1_HUMAN	Q15569 homo sapien
14	254.5	6.3	506	SRK4_SPOA	P42690 sporgilla 1
15	252	6.2	628	TESK_RAT	O63502 rattus norv
16	249.5	6.1	937	ROR1_MOUSE	O92139 mus musculu
17	245	6.0	1051	ULK1_MOUSE	O70405 mus musculu
18	244	6.0	937	ROR1_HUMAN	Q01973 homo sapien

19	243.5	6.0	505	1	FRK_HUMAN	P42685 homo sapien
20	243.5	6.0	815	1	MK07_HUMAN	Q13164 homo sapien
21	238.5	5.9	412	1	APRE_ARATH	P46573 arabidopsis
22	238.5	5.9	606	1	M3K7_HUMAN	O45318 homo sapien
23	238.5	5.9	1307	1	PHY1_CERPU	P25848 ceratodon p
24	238	5.9	805	1	FUSE_DROME	P23647 drosophila
25	236.5	5.8	497	1	SPKL_DUGRI	P42687 dugesia tlg
26	236	5.8	821	1	CIR1_ARATH	O05609 arabidopsis
27	235.5	5.8	323	1	FLK_RAT	P05760 rattus norv
28	235	5.8	505	1	SRK1_SPOA	P42686 sporgilla 1
29	233.5	5.8	806	1	MK07_MOUSE	O69578 mus musculu
30	232	5.7	579	1	PER_HUMAN	O20773 mus musculu
31	231.5	5.7	822	1	ST10_HUMAN	P16591 homo sapien
32	226.5	5.6	968	1	ST10_HUMAN	O94804 homo sapien
33	226.5	5.6	1055	1	FAK1_RAT	O35346 rattus norv
34	225.5	5.6	410	1	APRA_ARATH	O06548 arabidopsis
35	225	5.5	1052	1	FAK1_MOUSE	P23452 mus musculu
36	225	5.5	1453	1	CAL1_CHICK	P02457 gallus gall
37	224.5	5.5	943	1	ROR2_HUMAN	O01974 homo sapien
38	223.5	5.5	1223	1	K125_CABEL	Q10925 caenorhabd
39	223.5	5.5	1584	1	KYR1_DICDI	P18160 dictyostell
40	223	5.5	1182	1	ABR2_HUMAN	P42684 homo sapien
41	223	5.5	1460	1	CAL1_CANFA	O9xsj7 canis faml
42	222	5.5	507	1	ICK_CHICK	P42683 gallus gall
43	222	5.5	552	1	SRCL_DROME	P00528 drosophila
44	222	5.5	1520	1	ABL_DROME	P00522 drosophila
45	221.5	5.5	389	1	NAR_ARATH	P43293 arabidopsis

ALIGNMENTS

RESULT 1
M3KA_HUMAN
ID M3KA_HUMAN STANDARD; PRT; 954 AA.
AC 002779; Q12761; Q14871;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 10 (EC 2.7.1.37)
DE (Mixed lineage kinase 2) (Protein kinase MST).
GN MAP3K10 OR MLK2 OR MST.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96128179; PubMed=8536694;
RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
Sutherland G.R., Simpson R.J.;
RT "Complete nucleotide sequence, expression, and chromosomal
localisation of human mixed-lineage kinase 2.";
RT Eur. J. Biochem. 234:492-500(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95249256; PubMed=7731697;
RA Katoh M., Hirai M., Sugimura T., Terada M.;
RT "Cloning and characterization of MST, a novel (putative)
serine/threonine kinase with SH3 domain.";
RT Oncogene 10:1447-1451(1995).
RN [3]
RP SEQUENCE OF 244-480 FROM N.A.
RC TISSUE=Colon epithellum;
RX MEDLINE=93238756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT "Identification of a new family of human epithelial protein kinases
containing two leucine/isoleucine-zipper domains.";
RT Eur. J. Biochem. 213:701-710(1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC cerebrovascular disorders (thrombosis, vascular dementia); brain
CC disorders (phenylketonuria)); cancers (such as ovarian, lung, bladder,
CC liver, breast and lymphomas); behavioural disorders ('Outelet's
CC syndrome); musculo-skeletal disorders (arthritis, trauma, tendinitis);
CC renal disorders (nephrotic syndrome, glomerulonephritis); metastases of
CC malignancies and related disorders (leukemia, multiple myeloma); and
CC infections caused by bacteria, viruses, fungi and parasites. PTK is also
CC useful for screening therapeutic compounds. PTK is used as an antigen in
CC a vaccine to raise an immune response against infectious disease. PTK
CC nucleic acids are useful in gene therapy.

xx

Sequence 728 AA;

Alignment Scores:

Pred. No.:	3_73e-282	Length:	728
Score:	3822.00	Matches:	728
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	94.18%	Indels:	0
Dbl:	22	Gaps:	0

US-09-836-392-8_COPY_22-2205 (1-2184) x AAE00669 (1-728)

OY	1	ATGCGAGGGACCTCGGGGCCACCAGATGCATGAAGAAGATTCCGAGTCCGGCAGAG	60
Db	1	MetLeuAlaHisLysValArgLaIleThrAspAlaMetLysAsnPheserGIupheaGIinglu	20
OY	61	GCCAGCATGCTGCACGGSGCTGCACACCCCCTGCATCTGTGGCGCTCAATCGGCATCACATC	120
Db	21	AlaserMetLeuHnIALaleugHnIsProcySLetvalAlaleuilegilyleserile	40
OY	121	CACCCGCTCTGCTTCGGCCCTTGAGACTGGCGCCCTCAGACAGCTCAACACCGTGTCTCC	180
Db	41	HisproucycySphealAleugluleAlAproLeuSerSerleuasnThrValleuSer	60
OY	181	GAGAACGCCAAAGATTCTTCTTTATACCCCTGGGACACATGCTCACCCAAAATATGCC	240
Db	61	GLusnAlaAlaArgAspSerSerPheIleProleuIdlyHnImetLeuTrglInlysIIleaIA	80
OY	241	TACAGATCGCCTCGGGGCTGGCTACCTGCACACAAGAAAAACATCATCTTGTGACCTG	300
Db	81	TyrGlInleAlaserGIyleuAlaIatryLeuHnIsLysSnlleIlePhecySaBrleu	100
OY	301	AAGTCGCAACAATCTTGTTGTGTGCTCCCTTGACSTCAAGAGACATCAACATCAAGCTA	360
Db	101	LysSerAspAnlleLeuValITrrSerLeuAspAlaLySglInHnIsIleasnIleLysleu	120
OY	361	TCTACTAAGCGGATTGTGAGGCACATCATTCATGAGGGCCCCCTAGGCCCTGAGGGCACT	420
Db	121	SerAspyTrGIyleSerArGlnSerPheHnIsIdlyAlaleuGIylaIGluglyThr	140
OY	421	CGTGGCTACACGGCCCCAGAGATCAGAGCTCGCATTTATATATGATGGAAGTAGATATG	480
Db	141	ProglIyTrglInAlaProglInIleargProArGIlleValIlyTrspglInLysValasmet	160
OY	481	TTTCTATAGGAATGTGCTCTACGAGTCTCTGCAGACAGGCCCTGCACATGGGCCAC	540
Db	161	PheserTrglIymetValleuTyrglInleuLeuSerclYglInArGrroAlaleuGIynHnIs	180
OY	541	CACCAAGCTCCAGATTGCCAAGAACTGTCCAAGGGCATCCGGCCGGTTCGGGGCAGCCG	600
Db	181	HisGlnleuGlnIleAlaLysLysleuSerLysGIlyleArGrroValleuGIynGlnPro	200
OY	601	GAGGAAGCGAGTTCGGGGCAGCTCAGAGGCTCATATGATGAGNGTCGGGCACTAAGCCA	660
Db	201	GIUGlUVALGIInPheaArgIgleGlnAlaleuMetIetIducystrrPaPrInLysPro	220
OY	661	GAGAAAGCAGCCGCTGCGCTGTGCTGTGATAGCCAGATGAAGAAGACCGAATTTTGGCACC	720
Db	221	GIUlysaArgProleuAlaleuSerValIaSerGIImetLysAspProthrPheAlaItnr	240
OY	721	TTCAATGTAAGACTGTGCTGGGAGACAGACACCTTTCTTCATCCAGGGCCAGAG	780

Db 241 PHEMETYRGLILEUCYSCYSGLYSGINLTHRALAPHEPSESERGLNGLYGLN 260
 QY 781 TACACCGGGGTGTTGGATGAAAGAGAGTCCAGAACTACAGGTGTGTAACACA 840
 Db 261 TYRTHVALVALPHEPTRPSGLYSGINUSERRHARGASTTYRTHVALVALSNTHR 280
 QY 841 GAGAGGGGCTCATGAGAGTGCAGAGATGTCCTCCCTGGGATGAGGTGAGCTGCCAG 900
 Db 281 GLUYSGLYLEUMEGVALVALGINRMETCYSCYSPROGLYMETLYSVALSERCYSGIN 300
 QY 901 CTCGAGGTCCAGAGATCCCTGTGACAGCCAGAGACCAAAAATCTACATCTACACC 960
 Db 301 LEUGINVALGILARSERLEUTRPTHRALATHRGILUASPGILYSLERYITLEYRTHR 320
 QY 961 CTCAGGGGATGGCCCTTAACACACCCCAAGCGCTTGGAATCTCAGCTGCGTC 1020
 Db 321 LEULYSGLYMETCYSPROLEUSNTHRPROGLNHALALEUSPTHRPROVALVALVAL 340
 QY 1021 ACCCTGCTTGGCCGCTGCTGCTATTAAAAAGAAATCTTACCTGCTCTTACGGGCTC 1080
 Db 341 THRYSRPHLEVALVALPROVALIILEYLSYASNSERTYLEUVALLEUVALGLYLEU 360
 QY 1081 GCCGATGGCTGTGGCTGTGTTCCCGTGTCGGGGGACCCCAAGAGACAGCTCTCC 1140
 Db 361 ALAASPGLYLEUVALVALAVALPHEPROVALVALARGLYTHPROLYSASPSERCYSER 380
 QY 1141 TACCTGTGTCACACAGCAAGTCCCAAGTTCAGATCGCGGATGAGACGACGCG 1200
 Db 381 TYRLEUCYSETHISTRHAAASNRYSERLYSPHESETHLEAASPTLUSPALAARG 400
 QY 1201 CAGAACCCCTACCCAGTGAAGGCGATGAGGTGTCACAGCGGCTGAGTGTGTAC 1260
 Db 401 GLAASPROTYRPROVALYLSALAMETGLIVALAASNSERCYSERGILVALTRPTYR 420
 QY 1261 AGCAATGGGGGCGCTCTCTGTCATGACTGTGCTCCCTGGANATCTGACGGCGCTG 1320
 Db 421 SERASNGLYPROGLYLEUVALIILEASPCYSLASERLEUVALILECYSAARGAVAL 440
 QY 1321 GAGCCCTACATGGCCCCCTCCAGATGATGATGCTGCTGCTGAGTGCAGAGG 1380
 Db 441 GLUPRTYRMEALAPROSERMETVALTHRSERVALCYSESERGLUGLYARGGLY 460
 QY 1381 GAGGAGGTGCTGTGCTGCTGATGACAAAGCCCACTCTGTGTGATGATACACTCCAC 1440
 Db 461 GLUGLVALVALTRPCYSLEUASPAPLYLSALASNSERTYLEUVALMETYRHSERTHR 480
 QY 1441 ACCTACCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Db 481 THRTYRGLINLEUCYSLAARGTYRPHCYGLYVALPROSERPROLEUARGASPMETPHE 500
 QY 1501 CCGGTGGGGGCTTGGACAGCAAGCCCGGCGAGCCAGCAAGCGGCAACCCAAAGTG 1560
 Db 501 PROVALTRGPROLEUSPTHRGLUPROVALIILASERHISTRHAAASNPOLYSVAL 520
 QY 1561 CTTGAGGGGACTCATCCGCGAGCTGAGCATGATGATGAGAGAGTGGGCGAGCAG 1620
 Db 521 PROGLUGLYASPERIILEAASPAVALSERIILEMETYSERGILUGLYTHRGIN 540
 QY 1621 ATCTGTGACACCAAGAAATACACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
 Db 541 ILELEUITHESINGUSERTILEUTHRASPTLYCYSESERMETSERLYRYSERSETSER 560
 QY 1681 CCACCCGCGAGGCTGCGAGTCCCTCCCTCAAGCTCCCGCTCCCGAGAGAGTCTTCC 1740
 Db 561 PROPTORGLIILAAALAHASERPROSERSETSERLEUPROSERSETPROALASERSETSER 580
 QY 1741 AGTGTGCTTTCACACCACTGCGAGAGTCAACATGCTACATACGCGGCTGCTGCC 1800
 Db 581 SERVALTRPHESETHISTRHAAASPCYGLUASPSERAPMETLEUHSISTRPROGLYALVAL 600
 QY 1801 TCCGACAGGTCTGAGCATGACCTGACCCCATGAGCGGGAGACCTTCACGACGACTG 1860
 Db 601 SERASPTSRGERGILUNHSASPLEUTHRPTROMETASPGILYTHRPHESERGLNHISLEU 620

QY 1861 CAGCCGCTGAGACATCTCGCCCTCAGAGACTCATTTGGGTCCCAAGCGCGGTGAGAT 1920
 Db 621 GINALVALYSLILEUVALVALARGSPLEUITHETRPVALPROARGARGLYGLYASP 640
 QY 1921 GTTATTCGTCATTGGCCCTGAGAGAGATTTCTGAAGCCCGGGGCGGACTTGCCTCC 1980
 Db 641 VALIILEVALIILEGLYLEUGLILUSASPSERGLUALAGLNARGLYARGVALIILEVAL 660
 QY 1981 TTAAGAGCCCGAGACTGACTCGCATGGGAGTGTGCTGTGTAAGTGTGCCGCAAG 2040
 Db 661 LEULYSALARGILUENUTHRPROHSISGLYVALLEUVALASPAALVALVALALAYS 680
 QY 2041 GACACTGTGTGTGCACTTTGAAAAATGAACACAGAGTGTGCTGCTGCTGAGAG 2100
 Db 681 ASPTRHVALVALCYSTHRPHEGLUNSGILUASNTHRGLUTRPCYSLEUVALTRPARG 700
 QY 2101 GCGTGGGGCGCCAGGAGATTCACATTTTCTACAGTCTTACAGAGAGTGGCGGCTG 2160
 Db 701 GLYTRPGLYALARGGLUPHEASPILEPHETRYRGLNSERTYRGLINLEUGLIYARGLEU 720
 QY 2161 GAGGCTGCACCTCGCAAGAGAGG 2184
 Db 721 GLUALCYSTHRARGLYARGARG 728
 RESULT 2
 ABG64871
 ID ABG64871 standard; Protein; 706 AA.
 AC ABG64871;
 XX 27-AUG-2002 (first entry)
 DT
 XX
 DE Human albumin fusion protein #1546.
 XX
 KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cytosolic; antifertility; antiinflammatory; anticancer;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nocotropic;
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antiarthritic.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO20017137-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 12-APR-2001; 2001WO-US11988.
 XX
 PR 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI
 PI Rosen CA, Haseltine WA;
 XX
 DR WPI; 2002-010886/01.
 XX
 PT New fusion protein for treating disease e.g. diabetes comprises an
 PT albumin fused to a therapeutic protein.
 XX
 PS Claim 1: Page 1564-1566; 2102pp; English.
 XX
 CC The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA), also known as human serum
 CC albumin (HSA). The proteins are useful for treating a disease or
 CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its bioactivity.

Db 161 LeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlnProGluGln 180
 QY 607 GTGCAAGTCCGGGAGCTGCAGGGCGCTCATATGAGTGTGGGAGCACTAAAGCAGAGAG 666
 Db 181 ValGlnPheArgIleArgLeuGlnAlaLeuMetLeuLysTrpAspPheLysProGluLys 200
 QY 667 CGACCGGTGGCCCTGTGGGTGGAGCCAGATGAGAACCGGACTTTGGCACCTTATG 726
 Db 201 ArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThrPheMet 220
 QY 727 TATGAACTGTGCTGTGGGAGAGACAGCCCTTCTCATCCAGGCGCCAGAGATACAC 786
 Db 221 TygLeuLeuLysCysGlyLysGlnThrAlaPhePheSerSerGlnGlnGlnIleThr 240
 QY 787 GTGGTGTGGTGGAGTGAAGAAGAGAGCTCCAGAACTACACGGTGTGAACACAGAGAG 846
 Db 241 ValValPheThrPaspGlyLysGlnGlnLysArgAsnThrValValAlaThrGlnLys 260
 QY 847 GGGCTCATGAGGTGCAGAGAGATGTGCTGCTGGATGAGAGTGAAGTGGAGCTCCAG 906
 Db 261 GlyLeuMetGlnValGlnArgMetCysAspProGlyMetLysValSerCysGlnLeuGln 280
 QY 907 GTCCAGAGATCCCTGTGGAGAGACAGCCAGAGAGCAAGAAATCTACATCTACACCTCAG 966
 Db 281 ValGlnArgSerLeuThrPheAlaThrGlnAspGlnLysIleThrIleThrLeuLys 300
 QY 967 GGCATGTGCCCCCTAAACACAGCCCAACAGAGCTGTGATCTCCAGCTGTCACCTGC 1026
 Db 301 GlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspPheProAlaValAlaThrCys 320
 QY 1027 TTTTGGCCGTGCTGTATTAAGAAGATTCCTACTGCTGTAGGGGCTCGCCGAT 1086
 Db 321 PheLeuAlaValProValIleLysLysAsnSerThrLeuValLeuAlaGlyLeuAlaAsp 340
 QY 1087 GGGCTGTGTGCTGTGTTCCCGTGTGTGGGGAGACCCCAAGAGCAGCTGCTACTG 1146
 Db 341 GlyLeuValAlaValPheProValAlaArgGlyThrProLysAspSerCysSerThrLeu 360
 QY 1147 TGCTACACACAGCAACAGAGTCCAGTTCAGATCCGGATGAGAGCGAGCGAGAAC 1206
 Db 361 CysSerIleThrAlaAsnArgSerLysPheSerIleAlaAspGlnAlaArgGlnAsn 380
 QY 1207 CCCTACCCAGTGAAGGAGCATGAGAGTGTCAACAGCGCTGTGCTGTGTCAGCAAT 1266
 Db 381 ProTyrProValLysAlaMetGlnValAlaAsnSerLysSerGlnValIlePyrSerAsn 400
 QY 1267 GGGCCGGGCTCTTGTATGACTGTGCTCCCTGGAGATCTGACAGCGGCTGGAGACC 1326
 Db 401 GlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgArgLeuGlnPro 420
 QY 1327 TACATGGCCCCCTCATGTTAGTCACTGTGTGCTGAGCTGTGAGGGCAGAGGGAGAG 1386
 Db 421 TyrMetAlaIlePheSerMetValIleThrSerValLysSerSerGlnGlyArgGlnGln 440
 QY 1387 GTGTGTGTGGCTGTGATGAGCAAGGCCAACTCTGTGTGTGTACACATCCACCACTAC 1446
 Db 441 ValValIlePheCysLeuAspAspLysAlaAsnSerLeuValMetThrIleSerThrThrTyr 460
 QY 1447 CAGCTGTGTGCCCCGATCTTGTGCGGGGTCCAGGCCCTCAGGAGCAATGTTCCCGT 1506
 Db 461 GlnLeuCysAlaIleArgTyrPheCysGlyValProSerProLeuArgAspMetPheProVal 480
 QY 1507 CGGCCCTTGGACACAGAACCCCGGAGCCAGCACAGCGCCCAACCCAAAGTGCTGAG 1566
 Db 481 ArgProLeuAspThrIleProProAlaAlaSerIleThrAlaAsnProLysValProGln 500
 QY 1567 GGGGATTCATCGCGAGAGTGAAGCATCATATGAGTGAAGTGGGCGAGCGAGATCGT 1626
 Db 501 GlyAspSerIleAlaAspValSerIleMetTyrSerGlnLeuGlnIleThrGlnIleLeu 520
 QY 1627 ATCCACAGAGATCACTCACTGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
 Db 521 IleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSerProPro 540

QY 1687 CGCCAGGCTGCCAGGTCCCTCAAGCTTCCAGCTCCCCAGCAAGTCTTCCAGTGTG 1746
 Db 541 ArgGlnAlaIleAlaArgSerProSerSerLeuProSerSerProAlaSerSerVal 560
 QY 1747 CCTTTCGCCAGCAGTCCGCGAGAGCTCAGACATCTCATCTGCGCCGCTGCTCCGAG 1806
 Db 561 ProPheSerThrAspCysGlnAspSerAspMetLeuIleThrProGlyAlaAlaSerAsp 580
 QY 1807 AGCTGTGAGCATGACCTGACCCCATGAGAGCGGAGAGACCTTACAGCAGCACTGAGGCC 1866
 Db 581 ArgSerGlnIleAspLeuThrProMetAspGlyGlnThrPheSerGlnHisLeuGlnAla 600
 QY 1867 GTGAAGATCCTGCGCCGTGACAGACCTCATTTGGGTCCCGAGCGCGGTGAGATGTTATC 1926
 Db 601 ValLysIleLeuAlaValArgAspLeuIleThrValProArgArgGlyAspValIle 620
 QY 1927 GTCATGTGCTGGAGAGAGATTTCTGAAGCCACGCGGCGCGAGATCATGCGCTTAAAA 1986
 Db 621 ValIleGlyLeuGlnLysAspSerGlnAlaGlnArgGlyArgValIleAlaValLeuLys 640
 QY 1987 GCGCCAGAGCTGACCTCGCATGGGGGTGTGGATGCTGCGTGGGCGCAAGGACACT 2046
 Db 641 AlaArgGlnLeuThrProHisGlyValLeuValAspAlaValAlaLysAspThr 660
 QY 2047 GTTGTGTGCACTTTGAAATGAAGAACACAGAGTGTGTGCTGCGCGCTGGAGGGCTG 2106
 Db 661 ValValCysThrPheGlnAspGlnAsnThrGlnIlePheCysLeuAlaValIleArgGlyTyr 680
 QY 2107 GCGCCAGAGAGATTCACATTTTCTACAGTCTTACAGAGAGTGGCGCGCTGAGAGCT 2166
 Db 681 GlyAlaArgGlnPheAspIlePheThrGlnSerThrGlnLeuGlnIleArgLeuGlnAla 700
 QY 2167 TGCACCTGCAAGAGAGAG 2184
 Db 701 CysThrArgLysArgArg 706
 RESULT 4
 ABG64872
 ID ABG64872 standard; Protein; 706 AA.
 AC ABG64872;
 DT 27-AUG-2002 (first entry)
 DE Human albumin fusion protein #1547.
 KW Human albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW hematopoietic disorder; neural disorder; connective disorder;
 KW cytostatic; antifertility; antiinflammatory; anticancer;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nocotropic;
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antitarrtic.
 OS Homo sapiens.
 OS Synthetic.
 PN WO200177137-A1.
 PD 18-OCT-2001.
 PF 12-APR-2001; 2001WO-US11988.
 PR 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Haseltine WA;
 XX XX

DR MPI, 2002-010886/01.

XX New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein -

PS Claim 1; Page 1566-1569; 2102pp; English.

XX The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (Hb, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, reproductive disorders, digestive
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC (e.g. diabetes), hematopoietic disorders, neural disorders,
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
CC fusion proteins of the invention.

XX Sequence 706 AA:

Alignment Scores:

Pred. No.:	7,056-273	Length:	706
Score:	3700.00	Matches:	705
Percent Similarity:	99.86%	Conservative:	0
Best Local Similarity:	99.86%	Mismatches:	1
Query Match:	91.18%	Indels:	0
	23	Gaps:	0

US-09-836-392-8_COPY_22_2205 (1-2184) x ABG64872 (1-706)

QY 67 ATGTGTCACGGCTGTCACACCCCTGTCATCTGCGCTCATCGCATACATCCACCG 126
Db 1 MetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIleHisPro 20
QY 127 CTCGCTCGCGCCCTGGAGTGGCGCGCTGACAGCAGCTCAACACCGCTGTCGAGAAC 186
Db 21 LeucysPheAlaLeuGlnIleuAlaProLeuSerSerLeuAsnThrValLeuSerIuAsn 40
QY 187 GCCAGAGATTCCTCTTATACCCCTGGGACACATGCTCACCCAAATAATGCTTACAG 246
Db 41 AlaArgAspSerSerPheIleProLeuGlnHisMetLeuThrGlnIleAlaIleArgGln 60
QY 247 ATGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
Db 61 IleAlaSerGlyLeuAlaIleuAlaIleuHisIleuHisIleuHisIleuHisIleuHis 80
QY 307 GACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
Db 81 AsnAsnIleLeuValIleProSerLeuAspValIleGlnHisIleAsnIleLeuSerAsp 100
QY 367 TACGGGATTTTCAGAGCAGTATTCATGAGGCGGCTGAGGCTGAGGCGCTGCTGCG 426
Db 101 TyrGlyIleSerArgGlnSerPheHisGlnGlyAlaLeuGlyValGlnGlyThrProGly 120
QY 427 TACGAGGCGCCGAGATCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
Db 121 TyrGlnAlaProGlnIleuArgProArgIleValIleValIleValIleValIleValIle 140
QY 487 TATGGAATGGTCTGCTGAGTGTGCTGAGGACAGCGCTGCTGCTGCTGCTGCTGCTGCTG 546
Db 141 TyrGlyMetValLeuValIleuLeuSerGlyGlnArgProAlaLeuValIleHisGln 160
QY 547 CTCGAGATTCACAGAAAGCTGTCAGAGGCTACCGCGCTTCTGAGGCGAGCGAGGAA 606
Db 161 LeuGlnIleAlaIleValIleuSerIleuGlyIleArgProValLeuGlnIleProGlnGln 180
QY 607 GTGCAATTCGGCGAGCTGAGCGCTCATGATGAGAGTGTGAGACACTAAGCCAGGAAG 666
Db 181 ValGlnPheArgIleuGlnIleAlaLeuMetGluCysIleTrpAspThrIleuProGlnIle 200

QY 667 CGACCGTGGCCCTGTGCTGTGAGCAGATGAGACCCGACTTTTGGCACCCTTCATG 726
Db 201 ArgProLeuAlaLeuSerValIleValSerGlnMetIleuAspProThrPheAlaThrPheMet 220
QY 727 TATGACATGCTGCTGGAGACAGACGCTTCTTCATCCAGGGCGAGACTACAC 786
Db 221 TyrGlnLeuCysGlyIleGlnIleThrAlaPhePheSerSerGlnGlnIleuThr 240
QY 787 GTGCTGTTTGGATGGAAGAAGAGAGTCCAGACTACAGCTGGCTGGAACAGAGAG 846
Db 241 ValValPheTrpAspGlyIleGlnIleuSerArgAsnIleuValIleAsnThrGlnIle 260
QY 847 GGCCTCATGAGAGTGCAGAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
Db 261 GlyLeuMetGlyValGlnArgMetCysCysProGlnMetIleuValSerCysGlnLeuGln 280
QY 907 GTGAGAGATCCCTGTGGACACCCAGAGACCCAGAAATCTACATCTACACCTCAAG 966
Db 281 ValGlnArgSerIleuTrpThrAlaThrGlnIleuAspGlnIleuIleuIleuIleuIleu 300
QY 967 GGCATGTGCTGCTTAAACACCCCAAGCGCTTGTACTCCAGCTGTGCTGCTGCTGCTGCTG 1026
Db 301 GlyMetCysProLeuAsnThrProGlnIleAlaLeuAspThrProAlaValIleThrCys 320
QY 1027 TTTCTGCGCGCTGCTGCTTATTAAGAAATCTCTACCTGCTTGGCGGCTGCTGCTGCTG 1086
Db 321 PheLeuAlaValProValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 340
QY 1087 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146
Db 341 GlyLeuValAlaIleValPheProValIleValArgGlyThrProIleAspSerCysSerIleu 360
QY 1147 TGCTCAGACACAGCCAGAGGCTCAAGTTCAGATCGGAGGAGAGAGAGAGAGAGAGAGAG 1206
Db 361 CysSerHisIleThrAlaAsnArgSerIleuPheSerIleAlaAspGlnAspAlaArgGlnAsn 380
QY 1207 CCTACCCAGTGAAGCCCATGAGAGTGTGTCACAGCGGCTGTGAGCTGTGTCAGCAAT 1266
Db 381 ProIleProValIleAlaMetGlnValIleAsnSerGlySerGlnValIleTrpIleSerAsn 400
QY 1267 GGGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326
Db 401 GlyProGlyLeuLeuValIleAspCysAlaSerIleuGlnIleCysArgArgIleuIlePro 420
QY 1327 TACATGCGCCCTCCATGCTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386
Db 421 TyrMetAlaProSerMetValIleSerValIleCysSerSerGlnGlyArgGlyIleGln 440
QY 1387 GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
Db 441 ValValIleProCysLeuAspAspIleAsnSerIleuValMetIleuHisSerThrIleu 460
QY 1447 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
Db 461 GlnLeuCysAlaArgIleuPheCysGlyValIleProSerProLeuAlaArgAspMetPheProVal 480
QY 1507 GGGCGCTGAGACAGGAGACCCCGGAGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1566
Db 481 ArgProLeuAspThrGlnProPheAlaIleSerHisIleAlaAsnProIleuValIleProGln 500
QY 1567 GGGAGATCCATGCGGAGCTGAGATCATGTACAGTGAAGAGCTGAGGCGAGCATCTG 1626
Db 501 GlyAspSerIleAlaAspAlaSerIleMetIleuTrpSerGlnGlnIleuGlnIleu 520
QY 1627 ATCCACAGGAATCACTCACTGACTACTGCTCAATGCTCTCAATCCACACC 1686
Db 521 IleHisGlnIleuSerIleuThrAspIleuPheCysSerMetSerIleuSerSerProPro 540
QY 1687 CGCCAGGCTGCGAGTCCCGCTCAAGCTCCCGAGCTCCCGAGCAAGTCTTCCAGTGG 1746
Db 541 ArgGlnAlaAlaIleArgSerProSerIleuProSerProIleAlaSerSerIleuVal 560

XX Rosen CA, Komatsoulis GA, Baker KP, Bire CE, Soppet DR, Olsen HS:
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M,
P1 Ni J;
X
X WPI: 2002-304113/34.
DR N-PSDB; AAMD33695.
X
X An isolated nucleic acid molecule (1) comprising a polynucleotide which
PT encodes a polypeptide useful in the diagnosis and treatment of
PT disorders e.g., immune disorders -
X
X Claim 11; Page 472-474; 534pp; English.

AA033692-AAD33736 represent cDNAs corresponding to 21 human secreted protein genes, and AA021191-AA021235 represent the proteins they encode. AA021336-AA021280 represent human secreted protein fragments. The genes and their corresponding secreted proteins are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 21 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of immune or autoimmune diseases e.g., AIDS (acquired immune deficiency syndrome), asthma, anaemia and rheumatoid arthritis, breast neoplasia and breast cancer, neurological diseases e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, Tourette syndrome, meningitis, demyelinating disease, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, toxic neuropathies induced by neurotoxins, peripheral neuropathies, multiple sclerosis, leishmania and infarction, haemorrhages, schizophrenia, mania dementia, depression, panic disorder, learning disabilities, ALS, altered behaviours e.g., disorders in feeding, sleep patterns, balance and perception, encephalitis, disorders in cardiovascular, neural/sensory, reproductive and digestive systems, behavioural disorders and hyperproliferative disorder. The present sequence represents a human secreted protein of the invention.

SQ Sequence 706 AA;

Alignment Scores:

Pred. No.:	7,056-273	Length:	706
Percent Similarity:	3700.00	Matches:	705
Best Local Similarity:	99.86%	Conservative:	0
Query Match:	91.18%	Mismatches:	1
		Indels:	0
		Gaps:	0

US-09-836-392_8_COPY_22_2205 (1-2184) x AA021194 (1-706)

OY ATGTGTCAGGGCGTGCACACCCTTGATCGTGGCGCTCATCGGATCAGCATCCACCG 126
Db 1 MetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIleHisPro 20

OY CTCTGCTTCGCCCTTGAGACTGGCGCGCTCACAGACCTCAACACCGCTGTCGAGAAC 186
Db 21 LeucylSPeAlaLeuGlnIleAlaLaProLeuSerSerIleuAsnThrValLeuSerIuAsn 40

OY GCGAGAGATTCTTCCTTTATACCCCTTGAGACATGCTCACCCCAAAAATAGCCATACAG 246
Db 187 GCAGAGATTCCTTCCTTTATACCCCTTGAGACATGCTCACCCCAAAAATAGCCATACAG 246

OY AlaaArgApserSerPhelIleProLeuGlnHisMetLeuThrGlnIlyIleAlaTryGln 60
Db 41 AlaArgApserSerPhelIleProLeuGlnHisMetLeuThrGlnIlyIleAlaTryGln 60

OY ATTCGCTGGGCGCTTGCTACCTGCACAAGAAAAACATCATCTTCTGTGACCTGAAGTCG 306
Db 247 ATTCGCTGGGCGCTTGCTACCTGCACAAGAAAAACATCATCTTCTGTGACCTGAAGTCG 306

OY GACAACATTCGGTGGTGGTCCCTTGAGCTCAAGGACGACATCAACATCAAGTATCTGAC 366
Db 61 IleAlaSerGlyLeuAlaIlyrLeuHisIlySlysAsnIlellePheCysAAspleuYsser 80

OY AspasnIleLeuValITripSerLeuaspallysgIunHISileasnlIleLysleuSerasp 100
Db 81 AspasnIleLeuValITripSerLeuaspallysgIunHISileasnlIleLysleuSerasp 100

OY TACGGGATTTCGAGGCACTCATTCATCATAGGGCGCGCTTGAGCGTGCAGGCACTCTCTGC 426
Db 367 TACGGGATTTCGAGGCACTCATTCATCATAGGGCGCGCTTGAGCGTGCAGGCACTCTCTGC 426

Db 101 TyrGlyIleSerArgInSerPheHisGluGlyAlaLeuGlyValGluGlyThrProGly 120
 QY 427 TACCAGGCCCCAGAGATGAGGCTCGCATTTGATATGATGAGAGATGATGTTCTCC 486
 Db 121 TyrGlnAlaProGluIleArgProArgIleValIlyrAspGluLysValAspMetPheSer 140
 QY 487 TATGGAATGGTCTCTACGAGTGTCTGTCAGACAGCGCCCTGACATGGGCCACACAG 546
 Db 141 TyrGlyMetValLeuTyrGluLeuLeuSerGlyGlnArgProAlaLeuGlyHisHisGln 160
 QY 547 CTCGAGATTCGCAAGAGCTGTCAGAGGCAATCCCGGCTCGGGGAGCGGAGGAA 606
 Db 161 LeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGlnProGluGlu 180
 QY 607 GTGCAATTCGCGAGCTGAGCGGCTCATGATGAGTGTGTGGACACTAAGCCAGAGAG 666
 Db 181 ValGlnPheArgArgLeuGlnAlaLeuMetMetGluCysTyrPheThrLysProGluLys 200
 QY 667 CGACGCTGGCCCTGCTCGGTGTGAGCCAGATGAAAGACCCGACTTTTCCACCTTCATG 726
 Db 201 ArgProLeuAlaLeuSerValValSerGlnMetLysAspProThrPheAlaThrPheMet 220
 QY 727 TATGAACTGCTGTGGAGAGAGAGAGAGAGCTCTCTCTCATCCAGGGCCAGAGTACAC 786
 Db 221 TyrGluLeuCysCysGlyLysGlnThrAlaPhePheSerSerGlnGlyGlnGluTyrThr 240
 QY 787 GTGGTGTGGTGGAGTGAAGAAGAGAGTCCAGAACTACACGCTGTGTGACACAGAGAG 846
 Db 241 ValValPheThrPheProLysGluGlnSerArgAsnTyrThrValValAsnThrGluLys 260
 QY 847 GGCCTCATGAGAGTGCAGAGATGTCCTCCCTGGATGAGAGTGCAGCTCCAG 906
 Db 261 GlyLeuMetGluValGlnArgMetCysCysProGlyMetLysValSerCysGlnLeuGln 280
 QY 907 GTCCAGAGATCCCTGTGGAGAGAGAGAGAGAGAGAGAGATCTACACCTCCAG 966
 Db 281 ValGlnArgSerLeuThrPheAlaThrGlnLysGlnLysIleTyrIleTyrThrLeuLys 300
 QY 967 GGCATGTGCCCTTAAACACACCCCAACAGCGCTTGATCTCCAGCTGTCCACCTCG 1026
 Db 301 GlyMetCysProLeuAsnThrProGlnGlnAlaLeuAspThrProAlaValIlyrCys 320
 QY 1027 TTTCTTGGCCGTGCTTTTAAAGAAATTCCTACCTGTCTTACGGCCCTCGCGAT 1086
 Db 321 PheLeuAlaValProValIleLysLysAsnSerTyrLeuValLeuAlaGlyLeuAlaAsp 340
 QY 1087 GGGCTGTGGCTGTGTTCCCGTGGTGGGGGAGCCCAAGGAGACGCTGCTTACCTG 1146
 Db 341 GlyLeuValAlaValPheProValValArgGlyThrProLysAspSerCysSerTyrLeu 360
 QY 1147 TGCTCACACACCAACAGAGTCCAGATTCAGATCGCGATGAACGACGACGAGAAC 1206
 Db 361 CysSerHisThrAlaAsnArgSerLysPheSerIleAlaAspLysPheAlaArgGlnAsn 380
 QY 1207 CCTTACCCAGTGAAGCCATGAGGTGTGCACAGCGGCTGTGAGTCTGTGACAGCAAT 1266
 Db 381 ProTyrProValLysAlaMetGluValValAsnSerGlySerGluValIlyrPyrSerAsn 400
 QY 1267 GGGCCGGGCTCTTGTGATCAGCACTGTGCTCCCTGAGATTCGACGCGCGCGAGGCC 1326
 Db 401 GlyProGlyLeuLeuValIleAspCysAlaSerLeuGlnIleCysArgAlaGluGluPro 420
 QY 1327 TATATGGCCCCCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1386
 Db 421 TyrMetAlaProSerMetValIlyrSerValValCysSerSerLysGlyGlyGlnGlu 440
 QY 1387 GTGCTGTGCTGCTGCTGATGACAAGCCCAACTCTGTGTGTGATGATGATGATGATGAT 1446
 Db 441 ValValIlyrPyrSerLysAspSerLysAlaAsnSerLeuValIlyrMetTyrHisSerThrTyr 460
 QY 1447 CAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
 Db 461 GlnLeuLysAlaArgTyrPheCysGlyValProSerProLeuArgAspMetPheProVal 480

QY 1507 CGGCCCTTGACACGGAACCCCGGACCCAGCCACACAGCGCCCAACCCAAAGTGTCTGAG 1566
 Db 481 ArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAsnProLysValProLys 500
 QY 1567 GGGGCTCCATCCCGGAGGTAGATCATGATGATGATGATGATGATGATGATGATGATGATG 1626
 Db 501 GlyAspSerIleAlaAspValSerIleMetTyrSerGluGlnLeuGlyThrGlnIleLeu 520
 QY 1627 ATCCACAGGAATCACTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1686
 Db 521 IleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerTyrSerSerProPro 540
 QY 1687 CGCCAGGCTGCCAGGTCCCTCCCAAGCCTCCCGACCTCCCGACGCAAGTTCCTCCAGTGTG 1746
 Db 541 ArgGlnAlaAlaArgSerProSerSerLeuProSerProAlaSerSerSerSerVal 560
 QY 1747 CCTTCTCCACGACTGGAGAGACTCAGACATGATGATGATGATGATGATGATGATGATGATG 1806
 Db 561 ProPheSerThrAspCysGlnLysSerAspMetLeuHisThrProGlyAlaAlaSerAsp 580
 QY 1807 AGTCTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1866
 Db 581 ArgSerGlnHisAspLeuThrProMetAspGlyIlyrThrPheSerGlnHisLeuGlnAla 600
 QY 1867 GTGAGATCTCTGCGGTGACAGACCTCATTTGGGTCCCGAGCGGGGTGAGATGTTATC 1926
 Db 601 ValLysIleLeuAlaValAlaArgAspLeuIleTyrValProArgArgGlyLysAspValIle 620
 QY 1927 GTCATTTGGCTGGAGAGATCTGAGCCAGCGGCGGCGGAGTCTTCCGCTTAA 1986
 Db 621 ValIleGlyLeuGlnLysAspSerGlyAlaGlnArgGlyAlaValIleAlaValLeuLys 640
 QY 1987 GCCCGAGACTACTCCGATGGGT 2046
 Db 641 AlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValValAlaLysAspThr 660
 QY 2047 GTTGTGTGACCTTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2106
 Db 661 ValValCysThrPheGlnLysGlnLysAsnThrGluTyrPheLeuAlaValIlyrArgGlyTyr 680
 QY 2107 GGGCGGAGGAGTGCACATTTTCTACAGTCTTACAGAGAGTGTGGCGGCTGTGAGGCT 2166
 Db 681 GlyAlaArgGluPheAspIlePheTyrGlnSerTyrGluGlnLeuLysArgLeuGlnAla 700
 QY 2167 TGCATCTGCCAAGAGAG 2184
 Db 701 CysThrArgLysArgArg 706
 RESULT 6
 ID AAG67395 standard; Protein. 2014 AA.
 AC AAG67395;
 DT 13-NOV-2001 (first entry)
 DE Amino acid sequence of human protein kinase SGK258.
 KW Human; protein kinase; cancer; immune disease; cardiovascular disease;
 KW brain disease; neuronal disease; Alzheimer's disease; chromosome 15;
 KW Parkinson's disease; multiple sclerosis; metabolic disorder;
 KW peripheral nervous system disease; amyotrophic lateral sclerosis;
 KW infection; ocular disease; migraine; pain; sexual dysfunction;
 KW mood disorder; attention disorder; cognition disorder; hypotension;
 KW hypertension; psychotic disorder; dyskinesia; transplant rejection.
 OS Homo sapiens.
 PN WO200166594-A2.
 PD 13-SEP-2001.
 XX

QY 1501 CCGGTGGGGCCCTGGACAGGAGAACCCCGGACGACCCAGCCAGCCAGGCGCAACCAAGGTG 1500
 DB 1787 ProValArgProLeuAspThrGluProProAlaAlaSerHisThrAlaAspProLysVal 1806
 QY 1561 CCTGAGGAGGAGCTCCATCGCGAGCTGAGCATCATGTACAGTAGAGAGCTGGGACGAG 1620
 DB 1807 ProGluLysAspSerIleAlaAspValSerIleMetIleSerGluGluLeuGln 1826
 QY 1621 ATCTGATCCACAGAAATCACTCACTACTACTGCTCATGCTCTCTCTCTCTCTCTCTCT 1680
 DB 1827 IleLeuIleHisGlnIleSerLeuThrAspIleCysSerMetSerSerSerSer 1846
 QY 1681 CCACCCCGGCGGCTGGCAGGTCCCGCCAGCCCTCCACCCAGCCAGGAGTCTCTCC 1740
 DB 1847 ProProArgGlnAlaAlaAlaArgSerProSerSerLeuProSerProAlaSerSerSer 1866
 QY 1741 AGTGTGCTTCTCTCCACGAGCTGCGAGAGCTCAGACATGCTACATACGCGCGGTCTGCC 1800
 DB 1867 SerValProPheSerThrAspCysGluAspSerHisPheMetLeuHisThrProGluAlaAla 1886
 QY 1801 TCCGACAGGTCTGACATGACATGACCCCATGAGAGGGGAGACCTTCAGCCAGCACCTG 1860
 DB 1887 SerAspArgSerGluHisAspLeuThrProMetAspGluThrPheSerGlnHisLeu 1906
 QY 1861 CAGCGCGGAGAAATCTCTGCGCGCTCAGAGACCTCATTTGGGTCCCGGAGCGGCTGGAGAT 1920
 DB 1907 GlnAlaValLysIleLeuAlaValAlaArgAspLeuIleThrValProAlaArgGlyGlyAsp 1926
 QY 1921 GTTATGCTATGGCGCTGAGAGAGATCTGAGAGCCGAGCGGCGGATGCGCTC 1980
 DB 1927 ValIleValIleGlyLeuGluLysAspSerGlyAlaGlnArgGlyArgValIleAlaVal 1946
 QY 1981 TTTAAAGCCCGAGAGCTGACTCCGAGTGGGTGCTGTGTGATGCTGCCGTGGGCAAG 2040
 DB 1947 LeuLysAlaArgGluLeuThrProHisGlyValLeuValAspAlaAlaValAlaLys 1966
 QY 2041 GACACTGTGTGCTGACCTTTGAAAATGAAAACAGAGTGTGCTGCGCTCGGCGAGG 2100
 DB 1967 AspThrValValCysThrPheGluAsnGluAsnThrGluTrpCysLeuAlaValTrpArg 1986
 QY 2101 GCGTGGGCGCGCAGGAGATGACATTTTCTACAGTCTCTACAGAGAGAGCTGGCGGCTG 2160
 DB 1987 GlyTrpGlyAlaArgGluPheAspIlePheIleGlnSerIleGluGluLeuGln 2006
 QY 2161 GAGGCTTGCACCTCGCAGAGAGAG 2184
 DB 2007 GluAlaCysThrArgLysArgArg 2014
 RESULT 7
 AAE24137
 ID AAE24137 standard; Protein; 2014 AA.
 AC AAE24137;
 XX
 DT 23-SEP-2002 (first entry)
 XX
 DE Human kinase (PKIN)-8 protein.
 XX
 KW Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
 KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
 KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
 KW development; hepatitis; cardiovascular; hypertension; drug screening;
 KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
 KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
 KW hypercholesterolemia; obesity; gene therapy; cytostatic; anti-HIV;
 KW neuroprotective; hepatotropic; hypotensive; cardiant; nephrotropic;
 KW hyperlipidaemia; enzyme.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1280..1549

FT /note="Eukaryotic protein kinase domain"
 FT Domain 1295..1549
 FT /note-"Protein kinase domain"
 FT Domain 1367..1387
 FT /note-"Transmembrane domain"
 PN MO200233099-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 20-OCT-2001; 2001MO-US47728.
 XX
 PR 20-OCT-2000; 2000US-242410P.
 PR 27-OCT-2000; 2000US-244068P.
 PR 03-NOV-2000; 2000US-245708P.
 PR 09-NOV-2000; 2000US-247672P.
 PR 16-NOV-2000; 2000US-249565P.
 PR 22-NOV-2000; 2000US-252730P.
 PR 01-DEC-2000; 2000US-250807P.
 XX
 PA (INCYTE GENOMICS INC.
 XX
 PI Gururajan R, Baughn MR, Walla NK, Elliott VS, Xu Y, Arvizu C;
 PI Yao MG, Ramkumar J, Ding L, Tang YT, Hafalia AJA, Nguyen DB;
 PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM;
 PI Lai PG, Reclapon SA, Lu DAM, Borowsky ML, Thornton M, Swannaker A;
 PI Thangaveju K, Khan FA, Ison CH;
 XX
 DR WPI: 2002-454603/48.
 DR N-PSDB; AAD38851.
 XX
 PT New human kinase polypeptide, for diagnosing, preventing and treating
 PT cancer, immune system disorders, growth and development disorders,
 PT cardiovascular disorders and lipid disorders
 XX
 PS Claim 1; Page 147-152; 210pp; English.
 XX
 CC The invention relates human kinases (PKIN) and their corresponding
 CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
 CC treating and preventing cancer, an immune system disorder (e.g.,
 CC acquired immune deficiency syndrome (AIDS), Addison's disease, allergy,
 CC asthma, atherosclerosis, multiple sclerosis, psoriasis), disorders
 CC affecting growth and development (e.g., arteriosclerosis, cirrhosis,
 CC hepatitis), cardiovascular disorder (e.g., hypertension, myocardial
 CC infarction, Goodpasture's syndrome), and a lipid disorder (e.g., fatty
 CC liver, Gaucher's disease, Niemann-Pick's disease, hypercholesterolemia,
 CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
 CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
 CC condition or a disease associated with the expression of PKIN in a
 CC biological sample. A composition comprising PKIN or an agonist or
 CC antagonist of PKIN is useful for treating a disease or condition
 CC associated with decreased or increased expression of functional PKIN.
 CC PKIN is useful in a number of drug screening techniques and to analyse
 CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
 CC knockin humanized animals or transgenic animals to model human diseases,
 CC and in somatic or germ-line gene therapy. The present sequence is human
 CC PKIN protein.
 CC
 XX
 SQ Sequence 2014 AA;
 Alignment Scores:
 Pred. No.: 3,636-263 Length: 2014
 Score: 3574.50 Matches: 688
 Percent Similarity: 94.51% Conservative: 0
 Best Local Similarity: 94.51% Mismatches: 1
 Query Match: 88.09% Indels: 39
 DB: Gaps: 1
 US-09-836-392-8_COPY_22_2205 (1-2184) x AAE24137 (1-2014)
 QY 1 ATGCTGAGGACCTGCGGCGCAGCGATGAGAGAACTTCGAGTCCGCGAGAG 60
 DB 1326 MetLeuArgHisLeuArgAlaThrAspAlaMetLysAsnPheserGluPheArgGlnGlu 1345

OY	61	GCACAGATGCTGACAGGCTGTCAGACAGCCCTGATGCTGGCTATGGATACAGACT	120
Db	1346	AlaSerMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleIleIleSerIle	1365
OY	121	CACCGCTGTGCTTGCGCTGGAGCTGGCGCGCTGACAGCCTCAACACCTGTCTGC	180
Db	1366	HisProLeuCysProAlaLeuGlnIleValAlaProLeuSerSerLeuHisThrValLeuSer	1385
OY	181	GAGAACCGCAGAGATCTTCTTCCCTTATACCCCTGGGACACATGCTACCCCAAAATAGC	240
Db	1386	GluHisAlaArgAspSerSerPheIleProLeuGlnIleHisMetLeuHisThrIleHisIleAla	1405
OY	241	TACCAAGATCGCGCTGGCGCTGGGCTACCTGACACAAACAAATCATCTTCTGTGACCTG	300
Db	1406	TyrGlnIleAlaSerGluLeuAlaTyrLeuIleHisValGluHisIleIleIleHisValSerLeu	1425
OY	301	AAGTGGGACAAACATCTGTGGTGGTGGCTCCCTTGACCTGACAGAGACATCAATCAAGTA	360
Db	1426	LysSerAspHisIleLeuValIleTyrSerLeuAspValGluGlnHisIleAsnIleLysLeu	1445
OY	361	TCTGACTACGGGATTTTGGAGCGACATCTTCATACAGAGCGGCTTACGCGCTGGAGGCACT	420
Db	1446	SerAspTyrGluIleSerArgGlnSerPheHisGlnGluAlaLeuGlnTyrAlaIleGlnIleThr	1465
OY	421	CCCTGGCTACCAAGGCCCCACAGATCAGGCTCCGATTTGTATATGATGACAGGTAGATATG	480
Db	1466	ProIleTyrGlnAlaProGlnIleArgProIleArgIleValTyrAspIleLysValAspMet	1485
OY	481	TTTCTCATGGAATGGGTGCTCTACGAGTTGGCTGTACAGACAGCGGCTCCATCTGGCGCAC	540
Db	1486	PheSerTyrGlnMetValLeuTyrGlnLeuLeuSerGlnIleHisProAlaLeuGlnHis	1505
OY	541	CACCAAGCTCCAGATTGCCAAGAAAGTGTCCAAAGGCACTCGCGCGGATCTTGGCGAGCGG	600
Db	1506	HisGlnLeuGlnIleAlaTyrLysLeuSerTyrGluIleArgProValLeuGlnIlePro	1525
OY	601	GAGAGATGGAGTTTCCGGGACATGCAGAGGCGTCATATGATGATGCTGGGACACTAAGCA	660
Db	1526	GluIleValGlnPheArgAlaGluGlnAlaLeuMetGlnCysIleTyrAspIleHisPro	1545
OY	661	GAGAGAGCAGCGCTGGCGCTGCTGGTGGTGGACCAATACAGACCCGATCTTGGCCAC	720
Db	1546	GluLysArgProLeuAlaLeuSerValIleSerGlnMetLysAspProThrPheAlaThr	1565
OY	721	TTTCATGATGAATGCTGTCTGTGGAAAGCAGACAGCTTCTTGTATCCAGGGCCAGAG	780
Db	1566	PheMetTyrGlnLeuCysCysGluLysGlnThrAlaPhePheSerSerGlnIleGlnIle	1585
OY	781	TACACCTGGGTGTTTGGATGTGAAAGAGAGATCCAGAGATACACGGTGGTGAACCA	840
Db	1586	TyrThrValValPheIlePheArgGluLysGlnIleSerAlaHisTyrThrValAlaHisThr	1605
OY	841	GAGAAAGGCTCATGAGAGTGCAGAGAGATGTGCTGCCCTGGATGAAGTGAAGTGCAG	900
Db	1606	GluLysGluLeuMetGluValGlnArgMetCysCysProGluMetLysValSerCysGln	1625
OY	901	CTTCAGGATCCAGAGATCCCTGCTGGACAGCCACCGAGACCAAAAATCTACTCTACACC	960
Db	1626	LeuIleValGlnArgSerLeuThrPheAlaThrGln	1637
OY	961	CTCAAGGCAATGTGCCCTTAAACACACCCCAACAGGCTTGATCTCCAGCTGTCTGC	1020
Db	1637		1637
OY	1021	ACCTGCTTCTGGCCGTGCTGTATTAAAAAGATCTCTACCTGTCTTAAGCGGCTC	1080
Db	1638		1646
OY	1081	GCACATGGCTTGTGGCTGTGTTTCCCGTGGTGGCGGACCCCAAGACAGACGTCTGCC	1140
Db	1647	AlaAspTyrLeuValAlaValAlaPheProValValHisGluTyrProLysAspSerCysSer	1666

QY	1141	TACCTGGCTACACACAGCCACAGCTCCAAATTCAGATCGCCGAGAAAGACAGG	1200
Dp	1667	TYTCTCYSSTSRHTSHTRHAAASNAAGSETLSPHSESLIAIASPSLIHSPAIAR	1666
QY	1201	CAGAACCCCTTACCAGTAGAAGCCATGAGAGTGCTACACAGCGCTGAGAGCTGATC	1260
Dp	1687	GLIASNPROYTPROVALIYLSAIAIAMEGLIUALVALIASNSERGLYSERGLIUALIIPRY	1706
QY	1281	AGCAATGGGCGCGCCCTTGTTCATCGACTGTGGCTCCCTCGAGACATCGCAGCGGCTG	1320
Dp	1707	SERSNGLIYPROGLIYLEULAEVALIIEASPCYSAISERLEUGLIULEYSAIRGTARGL	1726
QY	1321	GAGCCCTACATGAGCCCTCATGTGTACGTCACTGCTGTGCAGCTGTGAGGCGAGAGG	1380
Dp	1727	GLUPROYMETALAPROSERMETVALIHPSERVALIYCYSESERGLUCLUYARGGLY	1746
QY	1381	GAGAGGCGCTGTGGTCCCTGGATGACAGAGCCAACTCTTGTGTATGTACACTTCAC	1440
Dp	1747	GLUCIUALVALITPCYSLAEUASAPBLYSALIASNSERLEUALIAMECTYRIASERH	1766
QY	1441	ACCTACACAGCTGTGTGCCCTGACTATCTGGGGGTCGCCAGCCCTCATGAGGACATGTT	1500
Dp	1767	THIRTYGLINUCYSAIAAAGTYRPHCYSGLYVALIPIROSERPROLEUADYASPHETHE	1786
QY	1501	CCCTGGGGCCCTTGGACACAGGAACCCCGGACGCCACACAGGCCAACCAAGTG	1560
Dp	1787	PROVALIARPROLEUASPTLIGUPROPROIALIASHTHIALASNPLOYVALI	1806
QY	1581	CCTAGAGGGAGCTACATCGCGGACGAGAGATCATGTACAGAGAGAGGTGGGACGAG	1620
Dp	1807	PROGLUCLYASPSRIIEALIASPVALSERIIEMETIYSERGLUGLIULEGLIYHICLI	1826
QY	1621	ATCTGTATCCACACAGATACTACTGATCTACTGCTCATGTCTCATGTCTCTATCCATCC	1680
Dp	1827	IIEULIENIISGLINCLISERLEUTHNPASPTYCYSESERMETSESTYIYSESERSE	1846
QY	1681	CCACCCCGGACGGTGCAGGTCCAGTCCCTCAAGCTCCCGAGTCCCGACAAAGTCTTCC	1740
Dp	1847	PROFROKARGIINALAIAAAGSETPROSESERLEUPROSESERPROLASERSE	1866
QY	1741	AGTGTGCTTTCTCCACCGACTCGAGAGACTCAGACATGCTACATACGCCGCTGCTCC	1800
Dp	1867	SERVALIPIROPHESERTHNPASPCYSGIULASPSERASPMETLEINHSTHNPROIYALIA	1886
QY	1801	TCCGACAGGTCTAGACATGACTGACCCCATGAGCGGGAGACTTCACCCAGCAGCTG	1860
Dp	1887	SERASPARISERTGIUNIASAPLEUTHNPROMETASPSLIYLIUTHPHSESTGIHISLEU	1906
QY	1861	CAGCCGCTGAAGATCCCTCGCGCTCAGAGACATATTGGTGTCCCGACGGCGGCTGAGAT	1920
Dp	1907	GLINAIVALIYSLIEULAIALVALIYASAPLEUIETTPVALIPIROHARGIYGLIYASP	1926
QY	1921	GTTATGTCATTGGCTGCGAGAAAGATTGTGAAGCCACGCGGGCGGAGTATTCCTC	1980
Dp	1927	VALIEVALIIEGLIYLEUGLIUYSASPSERGLYALAGIINARGIYARGVALIIEALVALI	1946
QY	1981	TTAAACCCCGGACAGCTGATCCGCATGGGGTCTGTGTGATGCTGCCGTGGTGGCAAG	2040
Dp	1947	LEUYSALAIARGIULEUTHNPROIHSGLIYVALIEUALIASPAILAIALVALIUALIAYS	1966
QY	2041	GACACTTGTGTGCACCTTGAATAATGAAACACACAGATGTGTGCTCGCTGCTGAGAG	2100
Dp	1967	ASPHRIYALVALCYSTHTRPHEGLIUAANGLIUAENTHGIUTPCYSLAUALIATIRPAR	1986
QY	2101	GGCTGGGGCGCCAGGAGTTGCAGATTTTCTACACAGTCCCTACGAGAGACTGGCGCGCTG	2160
Dp	1987	GLIYTPILYLAIRGILGULPHEASPIIEPHEITYGLINSEITYGLIUGLIULEGLIYARGLEU	2006
QY	2161	GAGGCTTGCACTGGCAAGAGAAAG	2184
Dp	2007	GLIUALCYSTHNPARGIYSAARG	2014

RESULT 8

ID ABG08051 standard: Protein; 809 AA.
 AC ABG08051;
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #8042.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensics;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI
 PI Dmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS72238.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 38410; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WPI
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SO Sequence 809 AA;
 Alignment Scores:
 Pred. No.: 3,41e-246 Length: 809
 Score: 3349.50 Matches: 651
 Percent Similarity: 98.49% Conservative: 1
 Best Local Similarity: 98.34% Mismatches: 7
 Query Match: 87.54% Indels: 4
 DB: 22 Gaps: 1
 05-09-836-392-8_COPY_22_2705 (1-2184) x ABG08051 (1-809)
 1 AUGCGAGGCGACGCGCGCCACCGAGCATGAGNACTTCGCGAGTCCGCGAGAG 60
 |||||||

Dd	149	MelleuAgnIstIsleuYgAlaThrAspAlaMetLysAsnPhSerGluPheArgGlnGlu	168
Oy	61	GCCAGCATGTGTGCACGGCTGCACAGCACCGCTCATCSTGGGGCTTCATCGGCATCAGATC	120
Dd	169	AlaserMetLeuHisAlaLeuGlnHisProCysIleValAlaLeuIleGlyIleSerIle	168
Oy	121	CACCGCGTGTGGTCCGCTGGAGCGCGGGCGCTCACAGACTCAACACCGTGCTGCC	180
Dd	189	HisProLeuYsPheAlaLeuGluLeuAlaProLeuSerSerLeuAsnThrValLeuSer	208
Oy	181	GAGAACCCSAGAAATTCTCTTTATACCSTGGGACACATGCTCTACCCAAAAATAGCC	240
Dd	209	GluAsnAlaArgspSerSerPheIleProLeuGlnHisMetLeuThGlnLysIleAla	228
Oy	241	TACCAATGCGCTCGGGCTGGGCTACSTGCACAAAGAAAACATCATCTTGTGCACCTG	300
Dd	229	TyrGlnIleAlaSerGlyLeuAlaTyrLeuHisLysLysAsnIleIlePheCysAspLeu	248
Oy	301	AAGCGGACAACATTTGTGGTGCGTCCCTTAGCGTCAAGGAGACATCAATCAATCACTA	360
Dd	249	LysSerAspAsnIleLeuValITrpSerLeuAspValLysGlnHisIleAsnIleLeuLeu	268
Oy	361	TCTSACTACGGGATTTTCGAAGCAGCATCATTCSCATGAAGGGCGCCCTAGCGGTGAGGCACT	420
Dd	269	SerAspTrpGlyIleSerArgInsérPheHisGlnGlyAlaIleGlyValaIGlGlyThr	288
Oy	421	CTTGCGCTACACGGCCCCAGAGATCAGGGCTCCCATTTGTATATGATGAAGAAGTACATATG	480
Dd	289	ProGlyTyrGlnAlaProGluIleArgProArgIleValTyrAspGluLysValAspMet	308
Oy	481	TTTCCTATGGAAATGGTGGCTCTCAGATGGTGTGTAGAGACAGCGGCTGCACCTGGCCAC	540
Dd	309	PheSerTrpGlyMetValaLeuTyrGlnLeuLeuSerGlyGlnArgProAlaLeuGlnHis	328
Oy	541	CACCACTCCAGATTTGGCAAGAAAGCTGTCCAAGGCGATCCGGCTTGSGGCACCGC	600
Dd	329	HisGlnLeuGlnIleAlaLysLysLeuSerLysGlyIleArgProValLeuGlyGlnPro	348
Oy	601	GAGAAAGTGCAGTTCCGGGCACATGCACAGGGCTCATATATGATGCTGGGACACTAAGCA	660
Dd	349	GlnGluValaGlnPheArgAlaLeuGlnAlaIleMetMetGlnCysTrpAspThrLysPro	368
Oy	661	GAGAAAGCGCCGCGGCTGTGGTGAGGAGACATTAAGA-CCGAGACTTTGGCCAC	719
Dd	369	GluLysArgProLeuAlaLeuSerValaSerGlnMetLysAlaArgAspPheCysHis	388
Oy	720	CTTCTATGTAAGTCTGTCTGTGGAGACAGACAGCCTTCTTCTATCCACAGGCCACAGA	779
Dd	389	LeuHisVal**ThrValaLeuTrpGlnAlaAspSerLeuLeuLeuIleProGlyProGly	408
Oy	780	GTAACCCGTGTGTTTTGGATGAGAAAAGAGAGCTCCAGAAACTACAGGGTGTAAACAC	839
Dd	409	ValHisAspGlyValaLeuGlnTyrLysArgGlyLysGlnGlnLeuHisIleGlyLysAsnThr	428
Oy	840	AGAGAAGGCGCTCATGAGAGGTCAGAGAGATGTCTCCCTGGGATGAAGTATAGCTGCCA	899
Dd	428	GlnLysGlyLeuMetGlnValaGlnArgMetCysCysProGlnMetLysValSerCysGcl	448
Oy	900	GCTCCAGCTCCAGAGATCCGTGTGGACAGCCACCGAGCAACGAAAAATCTACATCTACAC	959
Dd	448	AsnLeuIleValaGlnArgSerLeuITrpThrAlaThrGlnAspGlnLysIleTyrIleTyrTh	468
Oy	960	CCTCAAGGCGATGTGCCCCTTAAACACACCCCAACAGCGCTTGATATCCACAGCTGTGCl	1019
Dd	469	IleuLysGlyMetCysProLeuAsnITrProGlnIleAlaLeuAspThrProAlaValIla	488
Oy	1020	CACTGTCTTGTGCGCGCTGCTTATAAAGAAATTCCTACTGCTGTACGCGGCGCY	1079
Dd	488	IleHisPheLeuAlaValaProValIleLysLysAsnSerTyrLeuValaLeuAlaGlyLys	508
Oy	1080	CCGCATAGGGCTGTGGCTGTGTTTCCCGTGATGCGGGCGACCCCAAAGACAGCTGCTC	1139
Dd	508	UlaAspGlyLeuValaAlaValaPheProValaValaArgGlyThrProLysAspSerCysSe	528

QY 1140 CTACCTGTCTCAACACAGCCAGCTCCAGTTCAGCATGGCGGATGAAAGCCAGC 1199
 DB 528 rTyrLeuCySerHisThrIleAsnArgSerLysPheSerIleAlaAspGlnAspAlaAr 548
 QY 1200 GCAGAAACCCCTACCAAGTGAAGCCATGAGAGTGTCAACAGCGGCTCTAGAGCTGCTGTA 1259
 DB 548 gGlnAsnProTyrProValLysAlaMetGluValValAsnSerGlySerGluValAlaTrrPy 568
 QY 1260 CAGCAATGGGCGCGGCTCTTGTATCATGACTGTGCTCCCTGGAGATGTGACAGCGGCT 1319
 DB 568 rSerAsnGlyProGlyLeuLeuValIleAspCysAlaSerLeuGluIleCysArgArgyle 588
 QY 1320 GGAGCCCTATAGTGGCCCTCCATGGTTAGTCACTGTGTGAGCTGTGAGCTGTGAGGCGAAGC 1379
 DB 588 uGlnProTyrMetAlaProSerMetValThrSerValValCysSerSerGluGlyArgyl 608
 QY 1380 GGAGAGAGTGTGTGTGTGCTGATGATGACAGCCCAATCTCTGTGTATGATACCACTCCAC 1439
 DB 608 yGlnGluValValTrpCysLeuAspAspLysAlaAsnSerLeuValMetTyrHisSerTh 628
 QY 1440 CACCTTACCACTGTGTGTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1499
 DB 628 rThrTyrGlnLeuCysAlaArgTyrTrpCysGlyValProSerProLeuArgAspMetPh 648
 QY 1500 TCCCGTGGCGGCTGTGGACAGGACACCCCGGAGCCAGCCACAGCCCAACCAAGGT 1559
 DB 648 eProValArgProLeuAspThrGlnProProAlaAlaSerHisThrIleAsnProLysVa 668
 QY 1560 GCCTGAGGGGAGCTCCATCGCGAGCTGAGCATCATGTATGATGAGAGAGCTGGGAGCAGCA 1619
 DB 668 lProGlnGlyAspSerIleAlaAspValSerIleMetTyrSerGluIleuGlyThrGly 688
 QY 1620 GATCTGTATCCACCAAGATCACTCACTACTACTGTCTCATCTCTCTCTCTCTCTCTCTCT 1679
 DB 688 nIleLeuIleHisGlnGlnSerLeuThrAspTyrCysSerMetSerSerTyrSerSerSe 708
 QY 1680 CCCACCCCGGCGAGCTGCGCCAGTCCCGCTCAAGCTCCAGCTCCCGGACAGGTTCTTC 1739
 DB 708 rProProAlaGlnAlaAlaArgSerProSerSerLeuProSerSerProAlaSerSerSe 728
 QY 1740 CAGTGTGCTCTTCTCCACGAGCTGCGAGAGCTGAGACATGCTATACAGCCGCTGCTGC 1799
 DB 728 rSerValProPheSerThrAspLysGlnAspSerAspMetLeuHisThrProGlyAlaAl 748
 QY 1800 CTCGGAGAGCTGAGCATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1859
 DB 748 aSerAspArgSerGlnHisAspLeuThrProMetAspGlyGluThrPheSerGlnGlnLe 768
 QY 1860 GCAGGCC---GTGAGATCTCTCGCGCTCAGAGACTCATTTGGTCCCGGAGCGGCTGG 1916
 DB 768 uAlaGlyArgGluGln---SerSerProSerGluThrSerPheGlySerProGlyAlaValG 788
 QY 1917 AGATGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1976
 DB 788 lMetLeuSerSerLeuAlaTrpArgArgIleLeuAlaProSerGlyAlaGlnSerLeuP 808
 QY 1977 CG 1978
 DB 808 ro 808
 RESULT 9
 ABB61564
 ID ABB61564 standard; Protein; 2308 AA.
 AC ABB61564;
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 11484.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.
 OS
 XX
 PN WO200171042-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 PE
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR
 XX 11-JUL-2000; 2000US-0614150.
 XX
 XX
 PA (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX MPI; 2001-656860/75.
 DR
 XX N-PSDB; ABL05667.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions
 PS
 XX
 XX Disclosure: SEQ ID NO 11484; 21bp + Sequence Listing; English.
 CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
 CC sequences (ABLI01840-ABLI6175) and the encoded proteins
 CC (ABBI7737-ABBI2072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2308 AA:
 Alignment Scores:
 Pred. No.: 5,96e-19 Length: 2308
 Score: 364.50 Matches: 159
 Percent Similarity: 38.51% Conservative: 99
 Best Local Similarity: 23.73% Mismatches: 243
 Query Match: 8,988 Indels: 169
 DB: 22 Gaps: 27
 US-09-836-392-8_COPY_22_2205 (1-2184) x ABB61564 (1-2308)
 QY 4 CTGAGGCACTGTGGGCGCCAGCCATGGCCATGAAAGAACTTCTCCGAGTTCGGCGAGAGGCC 63
 DB 1684 LeuGlnHis-----SerCysLysAlaTyrCysThrAlaArgGlnGluLeu 1698
 QY 64 AGCATCTGCACGCGCTGTCAGACACCCCTGAGTGGCGCTCGGATCGCATGACATCCAC 123
 DB 1699 AlaValLeuLeuThrLeuLysHisProAsnIleValProLeuValGlyIleCysIleLys 1718
 QY 124 CCGCTGTGCTGTGCTGTGAGCTGCGCGCGCTCAGAGCCCTCAACACCGCTGTGCTGGAG 183
 DB 1719 ProLeuAlaLeuValLeuGlnLeuAlaProLeuGlyGlyLeuAspAlaLeuLysHis 1738
 QY 184 AACGCGAGATTCCTTCCTTTATACCCCTGGGACACATGCTACACCCAAATAAGCTTAC 243
 DB 1739 TyrArgArgSerGlyAlaHis-----MetGlyProHisThrPheGlnThrLeuValLeu 1756
 QY 244 CAGATGCGCTCGGCGCTGAGCTGACACAGAAATAATCATCTCTGTGACCTGAG 303
 DB 1757 GlnAlaAlaArgAlaIleGluTyrLeuHisArgArgArgIleIleTyrArgAspLeuLys 1776
 QY 304 TCGGACAAATCTGTGTGTGCTCT-----GACGTCAAGAG 342
 DB 1777 SerGlnAsnValLeuValTrpGluLeuProGlnProHisThrGlnAspSerProArgAsn 1796

```

QY 343 CACATCAACATCAAGCTATCTGACTACGGGATTTCGAGGCACTATTCATTCATGAGGCGCC 402
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1797 LeuValHisIleLeuValAspIleLeuValIleSerArgIleThrAlaProSerGlu 1816
QY 403 CTAGCGCGTGGAGGACATCCCTGGCTACACGAGCCGAGATCAGGCGCTCGCATTTATAT 462
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1817 LysGlyPheGlyGlyThrGluGlyPheMetAlaProGluIle-----IleArgTyr 1833
QY 463 GATGAGAGAGTAGATATATCTCTATATGAAATGGTGTCTCTAGAGTCTGTGACAGACAG 522
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1834 Asn----- 1834
QY 523 CGCCCTGCATGGGCGCCACACAGCTCCAGATTCGCAAGAAAGCTGTCCAGGCGATCCG 582
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1835 -----GlyHisGlu-----SerIleLysGluCysIleLeuGluGlySerArg 1848
QY 583 CCGGTTCTGGGGGACCGCGAGAAATGCATTCGCGCA-----CTCAGGCGCTCATGTATG 639
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1849 ProAlaLeuThrGln---ArgGluThrGlnPheProThrCysCysLeuAspLeuMetVal 1867
QY 640 GAGTCTGGGACACTAAGACAGAGAGAGGAGGAGCGGCTGTCTCGGTGAGCCAGATG 699
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1868 LeuCysTrpHisGluGlnProArgArgArgProThrAlaSerGlnIleValSerIleLeu 1887
QY 700 AAGGACCGGACTTTTGGCCACCTTCATG-----TATGAACTG 735
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1888 SerAlaProGluCysIleHisLeuAspValValAlaMetProHisSerGluIle 1907
QY 726 TGCCTGTGGAGAG---CAGACAGCTCTTCTTCATCCAGGCGGAGAGATACACCTGTGG 792
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1908 ValCysGlyValAlaPheGlnSerLeuValGlyMetGlyAspAspGlnArgCysGlyLeuGlu 1927
QY 793 TTTTGGGATGGAAGAGAGATCCAGAACTACAGCTGTGTGTAACAGAGAGAGGCG--- 849
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1928 LeuTrpLeuProSerPheGlySerArgIleAspIleLeuAspCysSerProSerGlySer 1947
QY 850 CTCATGAGAGAGTCAAGAGATGCTGCG----- 876
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1948 LeuLeuGlnCysAsnSerIleSerCysSerProGlnProGlnValAlaProProLysThr 1967
QY 877 -----CTGTGATGGAAGTG 891
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1968 ProGluAsnGlyAlaHisSerArgAlaArgSerAlaGlnArgLeuProLysMetAsnMet 1987
QY 892 AGCTGCAGAGCTCAGAGTCCAGAGATCCCTGTGTGACAGAGCCAGAGAGAGAGAAATCTAC 951
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1988 LeuCysCysCysLeuValGlyGluAlaIleTrpMetGlyAspValSerGlyAsnLeuHis 2007
QY 952 ATCTACACCCCTCAAGGCGATGTGCCCTTAAACACACACCCCAAGGCGCTGATCTCCA 1011
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2008 AlaTyrSerThrSerThrTyrAlaHisLeuPheSer-----TyrMetLeuAsp---Pro 2024
QY 1012 GCTGTGCTACCTGCTCTCTGTGCGCGCTGTATTAAGAAATTCCTCACTGCTCTTA 1071
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2025 AsnIleLysSerAlaValIleSerLeuValTyrMetGlnLysIleAla---ArgValAla 2043
QY 1072 GCGGCGCTGCGCGAGCTGTGTGCTGTGTCTCCGCTGCGGCGAGAGAGAGAGAGAG 1131
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2044 ValGlyThrHisAsnLys-----ArgValPheLeuValAspAlaThrGlnMetProSer 2061
QY 1132 AGCTGCTCTCTAC-----CTGTGCTGACAGACAGAGAGAG 1164
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2062 AsnCysAlaPheAlaGluLysSerPheValLeuThrGlnIleCysSerGlyPheValLeu 2081
QY 1165 AGCTCCAGATTGACATCCCGGATGAGAG----- 1196
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2082 HisAlaAlaCysSerValValValAspGlyIleTyrGluLeuTrpCysGlyGluIleAla 2101
QY 1195 GCAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2102 GlyLysIleHisValPheProLeuAsnGluAsnGlyValSerGlyHisGlnAlaLeuCys 2121
QY 1216 -----GTGAAGCGCATGAGAGTGCTGTCACAGCGCG 1245

```

```

Db 2122 HisSerGluIleProAsnLeuIleGluAspValLysValAlaAlaGlyMetCysSerAsnLeu 2141
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1246 TCTGAGCTGTGTGACAGAGATAGGCGCGGCGCTCTGTC----- 1284
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2142 SerHisValPheSerCysLeuTyrProIleCysMetValTyrGlnTrpAspValIleSer 2161
QY 1285 -----ATGAGCTGTGCTCTCTGAGATCTGC---AGCGGCTGGAG 1323
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2162 LysArgIleGluAsnLysLeuAspSerLysLeuLeuProCysSerGluSerLeuGln 2181
QY 1324 CCGTACATGGCGCGCTCATGTGTACGTGACGTGCTGTGCTGAGGCGAGAGAGAGAG 1383
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2182 SerIleAlaIleAspGlnLysValAsnLeuIleLysCys----- 2194
QY 1384 GAGTGTGTGTGTGCTGATGACAGAGCAAGCTCTGTGTATGTACACATCCAGAGAG 1443
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2195 -----GlnIleSerAlaLeuAlaHisAsnSerGluLeu 2206
QY 1444 TACAGCTGTGTGCGGCTACTTGTGCGGCGTCCGAGCGCGCTGAGGAGATGTTTCC 1503
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2207 Tyr---IleGlyThrThrTrpGlyCysLeuIleValAlaGluLeuHisThrLeuArgPro 2225
QY 1504 GTG-----CGGCGCTGTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1554
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2226 IleSerValPheArgProTyrGluAsnGluIleLysSerIleIleThrLeuSerLysAsp 2245
QY 1555 AAGTGCCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1614
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2246 AsnValPro-----LeuIleAlaThrIleGlyArgArgTyr---ArgSerLeuIle 2261
QY 1615 ACGCAGATCTGTATCCACAGAGATACATCACTAGCTACTGTCTCATGTCTCTCTACAG 1674
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2262 SerArgTyrValAspSerAlaGlnSerIleThr-----LysSerSerAlaVal 2277
QY 1675 TCATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1704
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2278 SerIleProThrHisGlyAlaAlaLysSer 2287
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 10
ID AAE22763
AC AAE22763;
DE 09-AUG-2002 (first entry)
XX Human mitogen activated protein kinase, MAP3K11.
XX Human, cytosolic; antisense gene therapy; screening; protein kinase;
XX cancer; liver; colon; tumour; inflammation; arthritic synovium; MAP3K11;
XX enzyme; mitogen activated protein kinase.
XX Homo sapiens.
XX WO200224947-A2.
XX 28-MAR-2002.
XX 20-SEP-2001; 2001WO-JB02237.
XX 20-SEP-2000; 2000US-233999P.
XX 02-OCT-2000; 2000US-237419P.
XX 02-OCT-2000; 2000US-237423P.
XX 04-OCT-2000; 2000US-238558P.
XX 10-MAY-2001; 2001US-290555P.
XX (KINE-) KINETEK PHARM INC.
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX Yoganathan T, Delaney AD;
XX WPI, 2002-394145/42.

```


QY 1856 ACCTGACGCCGTGAGATCTCTGCCG----- 1882
 DB 924 AlAcysThrPro-----ThrProThCysHisCysAlaGlyGlyLeuSerValLeu 940
 QY 1883 -----TCAGAGACCTCATTTGGGTGCCAGCGCGGTGAGAGACTTATTCG 1927
 DB 941 ThrPheHisProValThrAlaThrAlaProGlySerProAlaProGlyGlyLeuTrpGly 960
 QY 1928 TCATTTGCC-----TGAGAGAGG 1945
 DB 961 ThrAlaAlaGlyLeuTrpGlyTrpAlaGlnAlaThrGluThrValGlyAspTrpThrArg 980
 QY 1946 ATTCTGAGACCGAGCGCGGCGGCGCTCTTAAAGCCCGAGAGCTGATCCGC 2005
 DB 981 SerAlaValProValGlyArgGlyAlaSerArgAspProAlaProGlyAsp----- 997
 QY 2006 ATGGGGTGTGTGGTGGATGCTGCTGCTG----- 2035
 DB 998 -----GlySerProValTrpSerGlnLysCysGluLeuSerAlaThrGln 1012
 QY 2036 -----CAAGGACACGTGTGTGTGCACTTGAAGTGAAGAACACAGAGTGT--- 2083
 DB 1013 AlaProSerGlnLeuAspSerLeuProAlaThrValAlaTrpValGlyArgGlnAlaGlyLeu 1032
 QY 2084 -----GCTGTG-----CCGCTGTGAGAGCGGCTGCGGCCCA 2113
 DB 1033 AlaLeuLeuTrpLeuLeuProAsnProGlyHisGlyThrPro 1046
 RESULT 12
 AAE16259 standard; Protein: 656 AA.
 ID AAE16259
 AC AAE16259:
 XX 26-MAR-2002 (first entry)
 DT
 XX
 DE Human kinase PKIN-5 protein.
 XX
 KW Human; Kinase; PKIN-5; cancer; leukaemia; adenocarcinoma; osteoarthritis;
 KW Immune disorder; attherosclerosis; Crohn's disease; Hodgkin's disease;
 KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
 KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
 KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoporosis;
 KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
 KW Rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
 KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
 KW cardiovascular disease; hypertension; vasculitis; myocarditis; gout;
 KW congestive heart failure; ischaemic heart disease; lung tumour; obesity;
 KW fatty liver; Niemann-Pick's disease; gene therapy.
 KW
 KW
 OS Homo sapiens.
 XX
 XX
 FH Key
 FT 13..252 Location/Qualifiers
 FT /label- Protein_kinase_domain
 FT 14..197
 FT /label- Protein_kinase_domain
 FT 14..257
 FT /label- "Eukaryotic protein kinase domain"
 FT Domain
 FT /label- Protein_kinase_domain
 FT 14..253
 FT /label- Protein_kinase_domain
 PN W0200196547-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 14-JUN-2001; 2001WO-US19444.
 XX
 PR 15-JUN-2000; 2000US-212073P.
 PR 23-JUN-2000; 2000US-213467P.

PR 30-JUN-2000; 2000US-215651P.
 PR 07-JUL-2000; 2000US-216605P.
 PR 13-JUL-2000; 2000US-218372P.
 PR 25-AUG-2000; 2000US-228056P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Lai P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
 PI Gandhi AR, Tribouley CM, Walla NR, Yao MG, Lu DM, Greenwald SR;
 PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
 PI Baughn MR, He A, Thornton M, Hafalla A, Patterson C, Gurusarajan R;
 PI Lo TP, Khan F, Reardon SA, Azimzai Y, Policky JL, Ding L;
 PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
 XX
 DR WPI: 2002-090207/12.
 DR N-PSDB: AAD26452.
 XX
 PT New polypeptides, useful for diagnosing, treating or preventing
 PT disorders of growth and development, cardiovascular and lipid, and
 PT diseases such as cancer, comprise human kinase polypeptides .
 PS
 PS Claim 1; Page 143-144; 197pp; English.
 XX
 CC The invention relates to human kinase PKIN proteins and their
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for
 CC treating a disease or condition associated with decreased expression of
 CC PKIN and a composition comprising PKIN antagonist is useful for treating
 CC a disease or condition associated with overexpression of PKIN. The
 CC disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,
 CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease), immune disorder
 CC (acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
 CC attherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
 CC autoimmune thyroiditis, bronchitis, Crohn's disease, diabetes
 CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
 CC rheumatoid arthritis, Sjoren's syndrome, uveitis, ulcerative colitis,
 CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts), cardio
 CC vascular disease (arteriovenous fistula, hypertension, myocarditis,
 CC aneurysms, congestive heart failure, angina pectoris, avoculitis,
 CC ischaemic heart disease, chronic bronchitis, lung tumours), lipid
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
 CC hypochlosterolemia, obesity). PKIN DNA is useful for assessing
 CC toxicity of a test compound and in gene therapy. The present sequence
 CC is human PKIN-5 protein.
 XX
 SQ Sequence 656 AA;
 SQ
 Alignment Scores:
 Pred. No.: 2,46e-16 Length: 656
 Score: 327.50 Matches: 101
 Percent Similarity: 45.48% Conservative: 75
 Best Local Similarity: 26.10% Mismatches: 168
 Query Match: 8.07% Indels: 43
 DB: 23 Gaps: 13
 US-09-836-392-8_COPY_22_2205 (1-2184) x AAE16259 (1-656)
 QY 28 GCCATGAGAACTTCTCGGAG-----TTCGGCAGAGGCCGACGATG 69
 DB 33 A1AValLys11ePheAsnLysHis1ThrSerLeuArgLeuAlaArgGlnGluLeuValAl 52
 QY 70 CTGCAGCGGCTGCAGACACCCCTGCATCGTGCATGCGATCATGCAACCCGCTC 129
 DB 53 LeuGlyHisLeuHisHisProSerLeuLeuSerLeuLeuAlaAlaGlyLleArgProArg 72
 QY 130 TGCCTTGGCTGAGAGCTGCGCGGCTGCAGACGCTCAACACCCGCTGTCGAGAACGCC 189
 DB 73 MetLeuValMetGluLeuAlaSerLysGlySerLeuAspArgLeuGlnGlnAspLys 92
 QY 190 AGAGATTTCTTCTTATACCCCTGGGACACATGCTCACCCAAAATAGCCCTACAGATC 249


```

QY 190 AGAGATTCCTCTTATACCCCTGGGACACATGTCACCCCAAAAATAGCCTACAGATC 249
DB 479 Alaser-----LeuthrArgThrLeuGlnHisArgIleAlaLeuHisVal 493
QY 250 GCGTCGGCGCTGCTGCTACCTGCACAGAAAACATCTTCTGACCTGAGCTGCGAC 309
DB 494 AlaAspGlyLeuArgThrLeuHisSerAlaMetIleIleTyrArgAspLeuHisProHis 513
QY 310 AACATCTGGTGTGCTGCTTACCTGACGTCAGAGACACATCAACATCACTACTACTAC 369
DB 514 AaValIleuLeuPheThrLeuTyrProAsnAlaAlaIleIleAlaIleAlaAspTyr 533
QY 370 GGGATTTCGACGACATCATCATGAGGCGCCCTAGGCGCTGAGCGCTCTGCTGCTC 429
DB 534 GlyIleAlaGlnTyrCysAspArgMetGlyIleLysThrSerGlnGlyThrProGlyPhe 553
QY 430 CAGGCCCCAGAGATC---AGCCTCGCATTTGATATGATAGAGAGGTAGATATGTTCTCC 486
DB 554 ArgAlaProGlnValAlaArgGlyAsnValIleTyrAsnGlnGlnAlaAspValTyrSer 573
QY 487 ATGGAATGCTGCTCTACAGATGCTGCTG---TCAGGACAGCGCCCTGCTGCTGCTG 537
DB 574 PheGlyLeuLeuLeuTyrAspIleLeuThrThrGlyArgIleValGlnGlyLeuLys 593
QY 538 -----CACACACAGCTCCAGATTGCCAAGACGATGCTGCAAGGACATCCGCCCG 585
DB 594 PheProAsnGlnPheAspLeuGlnIleGlnGlyLysLeuProAspValLysGln 613
QY 586 GTTCTGGGCGACCGGAGAGATGTCAGTTCCGCGCATCGACGCTCATGATGAGTGC 645
DB 614 TyrGlyCysAlaPro-----TyrProMetValGlnLysLeuIleLysGlnCys 629
QY 646 TGGGACACTAAGCAAGACAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
DB 630 LeuLysGlnAsnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 649
QY 706 CCGACTTTGGCCACTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
DB 650 AlaGlnLeuValCysLeuThrArgArgIleLeuLeuProLysAsnValIleValGlnCys 669
QY 757 TTTCTCTCATCCAGGCGGACGAGATACCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 804
DB 670 MetValAlaThrHisAsnSerArgAsnAlaSerIleThrLeuLysGlnHisThr 689
QY 805 AAGAGAGAGTCCAGGACTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 861
DB 690 AspArgGlyGlnLeuSerPheLeuAspLeuAsnThrGlnGlnGlnGlnGlnGlnGln 709
QY 862 -----CAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 915
DB 710 AlaAspSerArgIleLeuLysLeuAlaLeu-----ValHisLeuProValGlnLys 726
QY 916 TCCCTGTG-----ACAGCCACCGAGGACAGACAAATCTACATCTACACCCCAAGGC 969
DB 727 GluSerTrpIleValSerGlyThrGlnSerGlyThrLeuLeuValIleAsnThrGlnLys 746
QY 970 ATGTCGCCCTTAACACACCCCAACAGGCTTGATGATCTGCTGCTGCTGCTGCTGCT 1029
DB 747 GlyLysLysArgHisThrLeuGlnLysMetThrAspSer-----ValThrCysLeu 763
QY 1030 TTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1083
DB 764 TyrCysAsnSerPheSerLysGlnSerLysGlnLysAsnPheLeuLeuValGlyThrAla 783
QY 1084 GATGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
DB 784 AspGlyLysLeuAlaIlePhe 790
RESULT 14
AAW87503
ID AAW87503 standard; Protein; 1212 AA.
XX

```

```

AC AAW87503;
XX
DE 23-FEB-1999 (first entry)
XX
DE Human N-methyl-D-aspartate receptor subunit encoded by clone NMDA22.
XX
DE Human N-methyl-D-aspartate receptor; NMDAR2C;
XX
DE NMDA-activated cation-selective ion channel; glutamate receptor.
XX
OS Homo sapiens.
XX
PN US5849895-A.
XX
PD 15-DEC-1998.
XX
PF 20-APR-1994; 94US-0231193.
XX
PR 20-APR-1994; 94US-0231193.
XX
PR 20-APR-1993; 93US-0052449.
XX
PA (SIBI-) SIBIA NEUROSCIENCES INC.
XX
PI Daggett LP, Lu C;
XX
DR WPI; 1999-069812/06.
XX
DR N-PSDB; AAW82909.
XX
PT DNA encoding N-methyl-D-aspartate receptor subunit - useful for the
PI assembly of functional glutamate receptor subunits
XX
XX Example 3; Columns 253-262; 203pp; English.
XX
CC The present sequence represents a human N-methyl-D-aspartate (NMDA)
CC receptor subunit (NMDAR). The nucleic acid sequence does not contain
CC the 366 5'-most nucleotides, by the insertion of 11 nucleotides between
CC nucleotides 1300 and 1301, nor the 15 nucleotides at positions 1300
CC 1960-1974, nor the 1061 3' nucleotides, as set forth in AAW82889. The
CC cDNA sequence is derived from clone NMDA21. The NMDAR subunits contribute
CC to the formation of NMDA-activated cation-selective ion channels. In
CC addition to being useful for the production of NMDA receptor subunit
CC proteins, the nucleic acids are also useful as probes to identify and
CC isolate nucleic acids encoding related receptor subunits. Functional
CC glutamate receptors can be assembled from several NMDA receptor subunit
CC proteins of one type (homomeric) or from combinations of subunit proteins
CC of different types (heteromeric). The present invention also comprises
CC methods for using such receptor subunits to identify and characterise
CC compounds which affect the function of such receptors, e.g. agonists,
CC antagonists and modulators of glutamate receptor function. The invention
CC also comprises methods for determining whether unknown protein(s) are
CC functional as NMDA receptor subunits.
XX
SQ Sequence 1212 AA;
XX
Alignment Scores:
Pred. No.: 1,36e-15 Length: 1212
Score: 319.00 Matches: 219
Percent Similarity: 33.25% Conservative: 64
Best Local Similarity: 25.73% Mismatches: 284
Query Match: 7,868 Indels: 284
DB: Gaps: 53
US-09-836-392-8_COPY_22_2205 (1-2184) x AAW87503 (1-1212)
QY 23 CCGATGCGATGAGAGACTTCTCGAGTCCGCGAGGAGGCGACACATCTGACCGCTGC 82
DB 477 ProSerProAspArgProSerTrpSerHisIleAlaLeuGlnCysGlyCysLeuSerCys 496
QY 83 AGC-----ACCCCTGCATCGTGGCGGCTCATCGGCA-----TCAGCATCC 121
DB 497 AlaSerLeuTrpTrpProSerProSerProSerSerCysSerSerThrSerAlaLeuSerAlaThr 516
QY 122 ACCGCG-----TCGTGCTTGCGCC----- 139
XX

```


Tue Apr 15 15:13:38 2003

us-09-836-392-8_22_2205.n2p.rag

Page 25

Oy 1977 CGTCF 1981
|||||
Db 942 roSer 943

Search completed: April 15, 2003, 07:07:23
Job time : 88.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 15, 2003, 06:31:27 ; Search time 33 Seconds
(without alignments)
12724.702 Million cell updates/sec

Title: US-09-836-392-8_COPY_22_2205
Perfect score: 4058
Sequence: 1 atctgcagcgacactgcgagc.....ctgcactgcagagagaag 2184

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame_n2p.model DEV=xlh
-O=/cgn2_1/USPTO.spool/US09836392/unat_08042003_090328_22799/app_query.fasta.1.2375
-DB=PIR-73 -QMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09836392_ECGN_1_1_38_@unat_08042003_090328_22799 -NCPU=6 -ICPU=3
-NO_XLPPY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOS -DEV=TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	403.5	9.9	1090	2 T33475	hypothetical prote
2	337.5	8.3	847	1 A53800	mixed-lineage prot
3	315	7.8	402	2 T10671	protein kinase hom
4	314	7.7	417	2 G84635	probable protein k
5	312.5	7.7	954	1 S68178	mixed-lineage prot
6	304	7.5	1091	2 JCS399	protein-tyrosine k
7	293.5	7.2	888	2 A53518	dual leucine zippe
8	290.5	7.2	888	2 A53518	serine/threonine p
9	280	6.9	1123	2 A39962	kinase-related tra
10	279.5	6.9	394	2 JU0229	mixed-lineage prot
11	278.5	6.9	546	1 D84535	probable protein k
12	273.5	6.7	981	1 F04VGM	gag-abl polypeptid
13	270.5	6.7	553	2 T04683	hypothetical prote
14	270.5	6.7	1171	2 T12956	hypothetical prote

15	270	6.7	1130	1 TVHDA	protein-tyrosine k
16	268.5	6.6	438	2 C86273	protein kinases ho
17	268	6.6	475	2 T12955	probable protein k
18	268	6.6	660	1 O0BE3	BHLFI protein - hu
19	267.5	6.6	1147	2 F86287	hypothetical prote
20	267	6.6	668	2 JC2363	protein kinase (EC
21	266	6.5	660	1 O0BE3	BHLFI protein - hu
22	263	6.5	1029	2 H86179	hypothetical prote
23	263	6.5	1338	2 T18287	protein-tyrosine k
24	261	6.4	1248	2 B96827	hypothetical prote
25	259.5	6.4	345	2 T05675	hypothetical prote
26	257	6.3	410	2 B35670	protein-tyrosine k
27	254.5	6.3	506	1 S24553	protein-tyrosine k
28	251	6.2	736	2 T05137	protein kinase hom
29	250	6.2	1106	2 T00405	hypothetical 119.5
30	249	6.1	406	2 T52626	probable mitogen-a
31	249	6.1	1015	2 T00726	probable serine/th
32	248.5	6.1	462	2 T02726	probable protein k
33	248.5	6.1	512	2 T49552	protein-tyrosine k
34	248.5	6.1	576	2 T36729	probable serine/th
35	247	6.1	1257	2 T00486	serine/threonine-s
36	245.5	6.0	392	2 T47684	serine/threonine-s
37	245	6.0	1051	1 J00051	serine/threonine-s
38	244	6.0	937	2 A45082	neurotrophic recep
39	243.5	6.0	505	2 T18396	protein-tyrosine k
40	243.5	6.0	815	2 B56708	extracellular sign
41	242.5	6.0	314	2 T16052	hypothetical prote
42	242.5	6.0	879	2 T02728	serine/threonine-s
43	242	6.0	1717	1 A45358	epidermal growth f
44	241	5.9	364	2 G71410	probable protein k
45	241	5.9	982	2 T06576	probable protein k

ALIGNMENTS

RESULT 1

T33475

hypothetical protein T27C10.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33475

R:Zhu, H.J.; Graves, T.; Hawkins, M.

submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid T27C10.

A:Reference number: Z21354

A:Accession: T33475

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1090 <ZHU>

A:Cross-references: EMBL:AF098504; PIDN:ANC67413.1; GSPDB:GN00019; CESP:T27C10.5

C:Genetics:

A:Gene: CESP:T27C10.5

A:Map position: 1

A:introns: 182/3; 260/2; 319/2; 796/3; 826/3; 979/3; 1038/2

Alignment Scores:

Pred. No.: 9.15e-13 Length: 1090
Score: 403.50 Matches: 163
Percent Similarity: 41.26% Conservative: 113
Best Local Similarity: 24.36% Mismatches: 277
Query Match: 9.94% Indels: 116
DB: 29 Gaps: 29

US-09-836-392-8_COPY_22_2205 (1-2184) x T33475 (1-1090)

OY	28	GCATGAGAACTCTCCAGTTCAGGAGCCAGCATCTGCAGCGTGACAGCAC	87
DB	457	AlaCysArgAlaTyrCysThrSerArgGlnGluLeuSerLeuSerArgMetLysHis	476
OY	88	CCCTGATCGTGGCGCTCATCGCATTCACATCCACCCGCTGCTCGCCCTGAGAGCTC	147
DB	477	ProAsnValIleGlyLeuValGlyValCysThrPheProLeuSerLeuValValGluLeu	496


```

OY 148 GCGCGCTCAGCAGCTCAACACCGTCTGTCCGAGAACGACGATTCCTTATATA 207
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 497 ALAPROLEUGLYALALEUHSNGLINLEUENGLYSERHISARGLYALAEGLYHRLYSLEU 516
OY 208 CCCCTGGGACATGCTCCACCAAAATAGCTTACAGATGCTCGGCGCTCGCTAC 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 517 SERLEUGLY-----VALLEUGLYSERIALVALGlnVALALAEARGALAEUGLYR 534
OY 268 CTCGACAGAAAACATCATCTTCTGTGACCTGAGTGGACACATCTGTGGTGGTCC 327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 535 LEUHSISERLAHISLEILETYRARGASPLEUYSERGLIUSNVALLEUGLYTRPARG 554
OY 338 CTGACGCT-----AAGAGACACATCAACATCAAGCTATCTGATACGCGATT 375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 555 PHEPROALAPROPHESERPROGLINHRASPRVALLEULEUYSLEUGLYSPRYGLYLE 574
OY 376 PCGAGGAGCTATTCATCAGC-----GGCGGCTAGGCGGTGAGGACCTCGCTACCG 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 575 SERARGSERVALLEUPROSERGLYGLYALALYSGLYPHEGLYLYRGLIUGLYRHEMET 594
OY 433 GCGCCAGAGATC-----AGGCGCTCATGTGTATGATGAGAGGTAGATATGTC 483
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 595 ALAPROGLIULEVALARGPHEASNGLYGLIUGLYUTYRTHRGlnLYSVALSPCYSPHE 614
OY 484 TCCTATGAGATGCTGCTCTACGAGTGTCTGTACGAGACAGCGCTCGACCTGGGCCACAC 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 615 SERPHEGLYMETPHELEUITYRGLIULEUHTHRLYUHPHEPROPHE-----GLUSER 632
OY 544 CAGCTCCAGATGTCACAAAGAGCTGTCCAGAGGATCCCGCGCTGTGGGCGACCGGAG 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 633 GLIUGLHNSVALYGLIUGLYMETLEUASPRGLYALALYPRVALLEUHPROHISGLI 652
OY 604 GAGTGCAGATTCGCGGAGCTCAGAGGCTCATGATGAGTGTGGGACACTAAGCCAGAG 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 653 LEULEUHPROTHRPROKMETLEUASPRLEUASPRLEUVALHISCYSTRPERSERLAHISPROGLI 672
OY 664 AAGCAGCGCTGCGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 673 SERARGPROSERSERSERGLINLEUVALGLYPHECYALALAPROGLIUSHE-----THRHS 691
OY 724 ATGTATGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 692 LEULEUASPRVALCYGLIUGLYMETLEUVALALAEUHPROTHRGLINLEUVALALAEGLY 711
OY 766 -----TCCAGGCGCAGAGAGTACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 810
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 712 ILETHRASPRGLIULEASPRASPRASPRASPRASPRASPRASPRASPRASPRASPRASPR 728
OY 811 GAGTCCAGACTACACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 867
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 729 SERGLYARGGLIUMETVALMETGLYCYSTRHGLINTGLYPHEVALASPRGLIUSYSER 748
OY 868 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 749 ILEGLIUEHPROHISARGLYLYSTYRVALSER-----LYSVALARGASPRSERVAL 765
OY 922 TGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 981
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 766 TRPSERCYASPRGLIUSGLYGLIINVALTHRVALTYR-----GLYLSERLEUHS 782
OY 982 AACACACCCCAACAGCCTTGATCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1041
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 783 GLIUTHGLYHIS-----LEUGLINLEUPROSERLEUASNGLYTHRLEULECYALAPRO 800
OY 1042 GTTATTAAAGAAATCTTACCTGCTTACGAGGCGCTGCGGAGGTGGTGGTGGTGGTGGTGGT 1101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 801 GLIUEHISERASHASPRVALLEULEU-----ILSERASPRYSGLINLEUVALLEU 818
OY 1102 TTTCCGCTGTGCGGGGCAACCAAGACAGCTGCTTACCTGTGTGTGTGTGTGTGTGTGTGTGT 1161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 819 LEUULYSLEU-----SERGLUSERASNSERVALSERHISLEU----- 830

```

```

OY 1162 AACAGCTCCAACTTCAGCATCGCGGATGAAAGACGACGAGAACCCCTACCGCTGAG 1221
DB 831 -----GLYTHRILEASPRSERPROTYRGLIUEARG 840
OY 1222 GCGATGAGAGTGTACACAGCGCTCT-----GAGCTGTGTGTACGAAATGGCGGGGC 1275
DB 841 THRALARHRLPHEUGLYASNGLYSERHARGGLINLEUVALGLYHISERGLIUE 860
OY 1276 CTCCTGTGCATC-----GACTGTGCTCCCTGAGATCTGACAGCGCTGAGCGCTTACATG 1332
DB 861 ARGILLESERLEHISHSISLEHISERASNASPRPHEASERPHESERSERLEUITYR 880
OY 1333 GCGCCCTCC-----ATGGTACGTCAGTGTGTGCACTGTGAGGCGACAGGGAG 1383
DB 881 LEUPROASPRASPRLYCYSLILEVALARGINLEUVALGLYSERLYASPRALA----- 897
OY 1384 GAGTGTGCTGTGCTGCTGTGATACAAAGCCAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1443
DB 898 GLIUSVALTRPILALAEUENGLIUSYSERLYSVALGlnMETVALGLIUALGLIUS 917
OY 1444 TACAGCTGTGTGCGCGTACTTCTGCGGGGTCCCGACGCGCTCGAGGACATGTTTCC 1503
DB 918 ARGGLINVALTHRGLYSERLEU-----ASPRILEARGLYSVALMETPRO 931
OY 1504 -----GTGCGGCGCTTGAGACAGCAACCCCGGCGACGACGACGACGACGACGACGAC 1551
DB 932 GLYSERGLIUTHRILEHISHSISLEHISMETGLIUMETALASERGLIASNLYRVALTHRCYS 951
OY 1552 CCAAAGTGTGCTGAG-----GGGAGCTCCATCGCGGAGCTGAGATCATGTATGAGT 1602
DB 952 ILEGLYLEUENGLIUALARGASNASPRGLYASPR-----GLIUEUITYRILEGLY 966
OY 1603 GAGGAGCTGGGCGCAGATCTCTGTATCCACAGAGATCATCTACTACTGTCTCATG 1662
DB 967 THRISERLYSGLYLEUENVALLEHISALATHRTHRLYU----- 980
OY 1663 TCCTCTACTCTCTATCCCGACCGCGGACGCTGCGAGGTCCCGCTCAAGCTCCCGACG 1722
DB 981 -----GLIUEUENVALTHRCYSARG----- 987
OY 1723 TCCCAGCAGACTTCTTCAGTGTGCTTCTTCACGAGCTGCGAGACTCAGACATGTGTA 1782
DB 988 -----PROPHIEGLIUGLYASPRILEHISERILECYSLIEU 999
OY 1783 CATACGCGCGGTGCTGCTCCGACAGGTCTGACATGACCTGACCCCATGAGCGGGAG 1842
DB 1000 GLIUGLUPROSERARGIUGLUGLUSNTHRARGGLYLSALATHRTHRLYUENSERHIALA 1019
OY 1843 ACCTTCAGCAGCAGCAGCTGTGAGCGCGTGAAGATCCGCGGTGAGAGAC-----TC 1893
DB 1020 SERSERGLUSERGLYLEUENGLYTRPVAL-ARGGLIUALYALSERGLIUTHRVALASPRARGPH 1039
OY 1894 ATTTGGTCCCGACGCGCGGTGAG 1918
DB 1039 EARGSERPROALATHRVALGLI 1047

RESULT 2
A:3800
Mixed-lineage protein kinase (EC 2.7.1.1) 3 - human
N:Alternative names: protein kinase PKL1, protein kinase SPRK
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A53800; 158395
R:Galio, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
J. Biol. Chem. 269, 15092-15100, 1994
A:Title: Identification and characterization of SPRK, a novel src-homology 3 domain-c
A:Reference number: A53800; PMID:94253068; PMID:8195146
A:Accession: A53800
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-847 <GAL>
A:Cross-references: GB:U07747; NID:9464027; PIDN:AAA19647.1; PID:9464028
R:ing, Y.L.; Leung, T.W.; Heng, H.H.; Tsui, L.C.; Lassam, N.J.

```



```

OY 1349 CGTACTCGTGTGACCTGTGAGGCGAG----- 1378
    |||
    |||
    |||
Db 694 Proleu1leCyspheserleu1ythrProaspSerProthrProAlaProleu1eu 733
OY 1379 -----GGGAGAGGTCCTGTGTCCTCGATGACAGAGCCACT 1417
Db 714 LeuaspLeu1ylleProval1GlyGlnArgSerAlaLysSerProArgArgLuglu 733
OY 1418 CCTGTGTGATGTACCACTCCACCACTACAGCTGTGTGCGGCTACTTCTGCGGGTCC 1477
Db 734 ProArgGlyGlyThrValSerProPro-----ProLysThrSerArgSerAla 749
OY 1478 CCAGCCCCCTCAGGAGCATGTTCCCGTGGCCCTTGACAGAGAACCCCGACACCA 1537
Db 750 ProGly1ThrProGly1Thr-----ProArgSerPro 759
OY 1538-----GCCACAGCGCCCAACCCAAAGGTGCTGAGGGGAGCTCCATGCGGAGC 1585
Db 760 ProLeuGlyLeu1leSerArgPro-----ArgProSerProLeu1ArgSer 774
OY 1586 TCAGCATCATGTACAGTACAGGAGCTGGGACAGCATCTGTATCCACAGAAATCACTCA 1645
Db 775 -----Arg1leasPro1Trp-----Ser 780
OY 1646 CTGACTACTGCTCCATGTCTCTCTACTCTCATCCACCCCGCCAGGCTCCAGTCCC 1705
Db 781 PheValSerAla-----GlyProArgProSerProLeu1ProSerPro 794
OY 1706 CCTCAAGCCTCCCGCTCCCGACAGATCTTCCAGTGTGCTTCTTCCACAGACGCG 1765
Db 795 -----GlnProAlaProAlaArgAla--Pro1ThrPhePhePro1aspSer 810
OY 1766 AGGACTCAGACATGTAACATACGCCCGGTCTGCTCCGACAGGTCTGACATGACCTGA 1825
Db 810 ----- 810
OY 1826 CCCCCTGAGACGGGAGAGCT---TCAGCCAGACCTGCGCGCCGTAAAGTCTGCGCC 1882
Db 811 ProPhe1TrpaspSerProProAlaAsn--ProPheGlnGlyProGlnaspCysArgAl 830
OY 1883 TCAGAGACATCATTTGGGTCCCGACGCGCGGTGG 1916
Db 830 ac1n1ThrLysaspMet1yAlaGlnAlaPro1Trp 841

RESULT 3
T10671
protein kinase homolog F6E21.90 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000
C/Accession: T10671
R/Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999
A/Reference number: Z16533
A/Accession: T10671
A/Molecule type: DNA
A/Residues: 1-412 <BEV>
A/Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90
A/Experimental source: cultivar Columbia; BAC clone F6E21
C/Genetics:
A:Gene: ATSP:F6E21.90
A:Map position: 4
A:Intons: 300/2
C/Superfamily: kinase-related transforming protein; protein kinase homology
F:135-392/Domain: protein kinase homology <KIN>

```

```

Alignment Scores:
Pted. No.: 2,14e-08 Length: 412
Score: 315.00 Matches: 82
Percent Similarity: 54.55% Conservative: 44
Best Local Similarity: 35.50% Mismatches: 73
Query Match: 7.76% Indels: 32
                Gaps: 9

```

```

US-09-836-392-8_COPY_22_2205 (1-2184) x T10671 (1-412)
OY 46 GAGTCCGGAGGAGGAGGACATGCTGACAGCGCTCGACAGCCCTGCATCGCGGCGTC 105
    |||
    |||
    |||
Db 176 GlnPheGlnGln1ValSerMetLeu1AlaPheLeu1y1s1Proasp1leVal1ArgPhe 195
OY 106 ATGGCATCATGATCCACCGGCTC-----TGCCTGCGCTGAGCTGCGCGCTCAGC 159
Db 196 IleGly1AlaCys1leLys1ProMetVal1TrpCys1leVal1ThrGln1ur1Arg1y1s1Gly 215
OY 160 AGCCTCAACACCGTGTCTGCTCCGAGAACGCCAGATCTCTTATACCTCTGGGACAC 219
Db 216 SerVal1Arg1n1Phe1Leu1Tr1y1s-----Arg1Gln1asn1rrg1Ala1Val1ProLeu1y1s1eu 233
OY 220 ATGCTCACCCAAATAATAGCCTACAGATCGCTCGGCGCTGAGCTTACCTGACAGAA 279
Db 234 AlaValMetGln-----AlaLeuaspVal1AlaArg1yMet1Ala1y1rVal1H1s1G1u1Arg 251
OY 280 AACATCATCTTCTGTGACCTGAAGTCGACACATTCGTGTGTGCTTGTACGTCAG 339
Db 252 Asn1Phe1le1H1s1Arg1asp1Leu1y1s1Ser1asp1asn1Leu1le-----Ser1Ala1asp----- 268
OY 340 GACACATCAACATCAAGCTATCTGACTAGGAGATTTCCAGGACATTCAT---GAG 396
Db 269 -----Arg1Ser1le1y1s1le1Ala1asp1Phe1y1Ala1Arg1y1le1G1u1Aln1Thr1G1u 286
OY 397 GCGCCCTAGCGCTGAGGAGCTCTGCTCGCTGAGGCGCCAGACAGATCAGCGCTCGCAT 456
Db 287 GlyMet1ThrProGln1ThrGly1Thr1y1r1Arg1y1r1Phe1Ala1ProGln1Met1le1Gln1H1s1Arg 306
OY 457 GTAATGATGAGAGAGTAGATATGTTCTCTATGAAATGCTGCTTACGATGCTGTCA 516
Db 307 Pro1y1r1Thr1Gln1y1s1Val1asp1Val1y1Ser1Phe1y1le1Val1leu1r1Pgl1u1le1Thr 326
OY 517 GGACAGCGCCCTGCTGAGGCGCCACACACAGCTCAGATGCGC---AAGAAGCTGTCCAG 573
Db 327 GlyLeu1eu1ProPheGln1asn1Met1H1Ala1Val1Gln1Ala1Phe1Ala1Val1asn1Arg 346
OY 574 GGCATCGCG-----CCGCTTCTGGGGGACCGCGGAGAGAGTG 609
Db 347 GlyVal1Arg1Pro1Thr1Val1Pro1Ala1asp1Cys1Leu1Pro1Val1leu1Gly1u----- 362
OY 610 CAGTTCGGGCGACTGACGGCGCTCATGATGAGTGTCTGGGACACTAAGCCAGAGACCA 669
Db 363 -----1leMet1H1r1Arg1Cys1Trp1asp1Ala1asp1ProGln1Val1Arg 375
OY 670 CCGCTGCGCTGTGCGTGTGAGCCACAGATGAG 702
Db 376 ProCysPheAlaGln1ur1le1Val1asn1Leu1Gln 386

RESULT 4
G84635
probable protein kinase [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C/Accession: G84635
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4420; M0ID:20083487; PMID:10617197
A/Accession: G84635
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-407 <STO>
A/Cross-references: GB:AE002093; NID:g4337195; PIDN:AAD18109.1; GSPDB:GN00139
C/Genetics:
A:Gene: At2g24360
A:Map position: 2
C/Superfamily: kinase-related transforming protein; protein kinase homology

```


C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
 F:110-448/Domain: protein kinase homology <KIN>
 F:118-196/Region: protein kinase ATP-binding motif

Alignment Scores:

Pred. No.:	6,896-08	Length:	1091
Score:	304.00	Matches:	202
Percent Similarity:	35.60%	Conservative:	96
Best Local Similarity:	24.13%	Mismatches:	272
Query Match:	7.49%	Indels:	267
DB:	2	Gaps:	44

US-09-836-392-8_COPY_22_2205 (1-2184) x S33596 (1-1091)

```

OY 13 CTGCGGGCCACCGATGCCATGAAGATCTCCGAGTCCGGCAGAGCCAGCATGCTG 72
DB 221 LeuSerGlnProGluAlaMet-----AspAspPheIleArgLysAlaMet 237
OY 73 CACGGCGTGCAGACCCCTGCATGCGGCTCATCGGCATCAGCATC-----CACCCGCTC 129
DB 238 HisSerLeuAspHisArgAsnLeuIleArgLeuTyrglyValValLeuThrProMet 257
OY 130 TCGTTCGCGCTGCAGCTGCGCGCTGCAGACGCTCAACACCGCTGTCCGAGAAGCC 189
DB 258 LysMetValThrGluLeuAlaProLeuGlySerLeuLeuAspArgLeuArgLysHisGln 277
OY 190 AGAGATTCTCTTATACCCCTGGGACACATGCTC-----ACCCAAATAATAGCC 240
DB 278 -----GlyHisPheLeuGlyThrLeuSerArgTyrAla 289
OY 241 TACAGATCGCTCGGCGCTGCAGCTACCTGCACAGAAAACATCATCTTCTGTACCTG 300
DB 290 ValGlnValAlaGlnGlyMetGlyTyrLeuGlySerLysArgPheIleHisArgAspLeu 309
OY 301 AAGTCGACACATCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 310 AlaAlaArgAsnLeuLeuAlaThrArgAspLeu-----ValLysIle 324
OY 361 TGTGACTGCGGATTTGAGG-----CAGTCATTCATGAGGGCGCCCTAGGC 408
DB 325 GlyAspPheGlyLeuMetArgAlaLeuProGlnAsnAspHisLysTyrValMetGlnGly 344
OY 409 GTGAGGCGCATCTCT-----GCTACGACGCCGACAGATCAGGCTCGATGTATATGAT 465
DB 345 HisArgLysValProPheAlaTrpCysAlaProGlnSerLeuLysThrArgThrPheSer 364
OY 466 GAGAGGTGATATGTCTTCTTATGAGATGAGTGTCTACGAGTGTGCTCA---GGACAG 522
DB 365 HisAlaSerAspThrTrpMetPheGlyValThrLeuTrpGluMetPheThrTyrGlyGln 384
OY 523 CGCCGCGACCTGGGCGACACGACGCTCAGATGCCAGATGCCAGTCCAG---GGCATC 579
DB 385 GluProTrpIleGlyLeuAsnGlySerGlnIleLeuHisLysLysAspLysGluGlyGln 404
OY 580 CGCCCGGTTCTGGGCGCGGAGAGATGCACTCCGAGTCCGAGCGCTCATGATG 639
DB 405 Arg-----LeuProArgProGlnAspCysPro---GlnAspIleTyrAsnValMetVal 421
OY 640 GAGTGTGAGACACTAAGCCAGAGAGACGACGCTGCGCTGTGCTGTGAGCCAGATG 699
DB 422 GlnCysTrpAlaHisLysProGlnAspArg----- 431
OY 700 AAGGACCCGACTTTGCAACCTTCATGATGAGTGTGCTGTGAGAGCAGACAGCTTC 759
DB 432 -----ProThrPheValAlaLeuArgAspPheLeuLeuGlnAlaGlnProThrAspMet 449
OY 760 TGTCTATCCAGGCGAGAG-----TAC 793
DB 450 ArgAlaLeuAlaHisPheGluGlnProAspLysLeuHisLysLysMetAsnAspValIle 469
OY 784 ACCGTGTGTGTGGATGAGAAAGAGAGCTCAGAGAACTACAGCGGTGTGAACACAGAG 843
DB 470 ThrValIle-----GluGlyArgAlaGlnAsnTyrTrpArgGlyGlnAsnThr--- 486

```

```

OY 844 AAGGCGCTCATGAGAGTGCAGAGGATGTGCTCCCTGGG-----ATGAG 888
DB 487 -----ArgThrLeuCysValGlyProPheProArgAsnValVal 499
OY 889 GTGAGCTGCCAGCTCCAGGCTCCAGAGATCCCTGTGGACAGCCACCGAG----- 937
DB 500 ThrSer-ValAlaGlyLeuSerAlaGlnAspIleSerGlnProLeuGlnAsnSerPheI 519
OY 938 ---ACCGAAATCTACATCTACACCTCCAGGCGATGTCGCCCTTAAACACCCCA 993
DB 519 enIleThrGlyHisGlyAspSerAspPro---ArgHisCysTrp-----GlyPheProAs 536
OY 994 CAGGCGCTGATCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1053
DB 536 PargIleAspLeuLeuTyrglyLeuGlyAsnProMetAspProProAspLeuLeuSerValG 556
OY 1054 AAT----- 1071
DB 556 uLeuSerThrSerArgProProGlnHisLeuGlyGlyValLysLysProThrTyrAspPr 576
OY 1072 GCGGCGCTCG----- 1081
DB 576 ovalSerGlnAspGlnAspProLeuSerSerAspPheLysArgLeuGlyLeuArgLysPr 596
OY 1082 -----CCGATGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1134
DB 596 oGlyLeuProArgGlyLeuTrpLeuAlaLysProSerAlaArgValProGlyThrLysAl 616
OY 1135 TGT----- 1149
DB 616 AsenArgGlySerGlyAlaGlyValThrLeuLeuAspPheGlyGlnGluProVal----- 634
OY 1150 TCACACAGCGGCAACAGGCTCCAGTCCAGATCGCGGATGAAGAGCAGCAGCAAGCC 1209
DB 635 ---ValProProTyrGlyProAlaAlaLeuProGlyAlaAlaGlnHisGlyArgLeu-L 653
OY 1210 TACCC-----AGTGAAGCCATGAGAGGCTGCAACAGCGCTGAG 1251
DB 653 euProAlaGlyArgAspProAlaSerGlnProHisAlaGlyThrAlaArgProLeuHisP 673
OY 1252 GTCTG---GTACAGCAATGGCGGG-----CCTCCTGTCTAT 1286
DB 673 roThrProValAlaSprTrpAspAlaArgProLeuProProProAlaTyrAspAspV 693
OY 1287-----CGACTGTG---CTCCCTGAGATCTGCAGGCGCTG 1320
DB 693 alAlaGlnAspLeuAspAspPheGlnIleCysSerIleAsnSerThrLeuValGlyAlaG 713
OY 1321 GAGCCTACATGCGCCCTCCATGTTACGTACGTACGTGCTGCACTGTGAGGCGAG--- 1376
DB 713 LysValProAlaGlyProSerGlnGlyGlnThrAsnTyrAlaPheValProGlnAlaAla 733
OY 1377-----AGGAGAGAGCTCGCTGTG 1395
DB 733 rGProProProProLeuGlnAspAsnLeuPheLeuProGlnGlnIleGlyLysLys--- 751
OY 1396 TGGCTGATGACAGAGCAACTCCTGTGATGATACACTCCACACCTACAGAGCTGTGT 1455
DB 752 --Pro-ProSerSerAlaIleThrAlaGlnIlePheGlnHisLeuGlnGlnGlnCysne 770
OY 1456 G-----CCGCTACTTCTGCGGCGTCCAGCGCCCTCAGGAGCATGTTT 1500
DB 770 tArgGlnLeuGlnAlaProGlySerProAlaProSerProGlnGlyLysAspArgLys 790
OY 1501 CCGCTGCGCTGTTGGACAGGAGACGCCGCGAGCTCAGCCACAGCCAACTCAAGGGT 1500
DB 790 sProGlnValPro-----ProArgValPro----- 798
OY 1561 CCGTGAAGGGAGCTCCAGCGGAGCGTGAAGCATCATGTACAGTAGAGAGTGGCAGCGAG 1620
DB 799 -----LleProProArg----- 802

```

QY 1621 ATCCGATCCACGAGATCACTGACTGCTCCATGTCCTCT----- 1669
 DB 803 -----ProthArgProHis---ValGlnLeuSerProAlaProProGlyGluGlu 819
 QY 1670 ----ACTCTCATCCCGCCGCGCCAGGTGCCAGGTCCCGCCCAAGCCGCCAGCTCC 1725
 DB 819 uhrSerGlnTPProGlyProAlaSerProProAlaGlyProProAlaGluPro----- 837
 QY 1726 CCAGCAAGTCTTCACAGTGTGCTTCTCCACCGAGCTGAGAGCTACAGATGACAT 1785
 DB 838 -----LeuSerProGlnGlySerArgThProSer----- 847
 QY 1786 AGCCCGGCTGCTCCCGCCAGAGCTGTGACATGACCTGACCCCAAGAGCGGAGACC 1845
 DB 848 ----ProLeuValProProGlySerSerProLeu-----ArgLe 861
 QY 1846 TTCAGCAGCAGCTGACGAGCGGTGAGAGTACCTGCGTCCAGAGCTTCATTTGGTCCC 1905
 DB 861 uSerSerSer-----ProGlyLysThrMetProThrGlnSerPheAlaSer-A 878
 QY 1906 AGCGCGGTGAGAGATGTATCGTCAATGGCTGAGAGAGATTCGAGGCCGCGGCG 1965
 DB 878 sProLysTyr----- 881
 QY 1966 CGAGTATATGCGCTTTAAAGCCGAGAGCTGATCCGATGGGCTGCTGTGATGCT 2025
 DB 882 -----AlaThrProGlnValIleGlnIleAlaProGlyAlaGlyGlyProC 896
 QY 2026 GC-----CGTGTGCGCAAGAGACACTGTGTGTGACACCTTGAATAATGA 2070
 DB 896 ysIleLeuProIleValArgAspGlyLysLysValSerSerThrHis----- 911
 QY 2071 AACACAGACTGTGCTGCGCTGCGCTGTGAGAGGCGCTGGGCGC 2111
 DB 912 --TyrTyrLeuLeuProGlnArgProSerTyrLeuGlnArg 924

RESULT 7

JC5399

dual leucine zipper kinase (EC 2.7.-.-) - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 28-May-1999
 C:Accession: JC5399

R:Matsumi, N.; Sarkar, G.; Shuto, T.; Morris, J.; Bronx, J.T.; Mizuno, K.; Bolander, M.E.
 Blochm. Biophys. Res. Commun. 229, 571-576, 1996

A:Title: Identification of a dual leucine zipper kinase involved in rat fracture repair.
 A:Reference number: JC5399; PMID:97127443; PMID:8954939

A:Accession: JC5399
 A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA
 A:Residues: 1-888 <MAT>

C:Comment: This enzyme is involved in regulating cell function in the musculoskeletal system.
 C:Genetics:

A:Gene: IDLK
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C:Keywords: ATP; phosphotransferase
 F:54-75, 89-98/Region: glycine-rich

F:156-454/Domain: kinase catalytic #status predicted <CAT>
 F:156-404/Domain: protein kinase homology <KIN>

F:164-172/Region: protein kinase ATP-binding motif
 F:421-449/Region: leucine zipper motif

F:472-500/Region: leucine zipper motif
 F:557-888/Region: glycine-serine-proline rich #status predicted

Alignment Scores:
 Pred. No.: 2, 29e-07 Length: 888
 Score: 293.50 Matches: 179
 Percent Similarity: 33.38 Conservative: 78
 Best Local Similarity: 23.25 Mismatches: 256
 Query Match: 7, 238 Indels: 257
 DB: 2 Gaps: 33

US-09-836-392-8_COPY_22_2205 (1-184) x JC5399 (1-888)

QY 28 GCCATGAGAACTTCTCCAGTTCGGGCGAG---GAGGCCAGCATCTGACGCGCTGCGAG 84
 DB 183 AlaValIleGlyValAlaArgAspLeuGlyThrAspIleGlnHisIleuArgLysLeuLys 202
 QY 85 CACCCCTGATGCTGGCGCTCATCGGCATCAGCATCCAGCATCCAGCCGCTGCTGCC--- 138
 DB 203 HisProAsnIleIleThrPheLysGlyValCysThrGlnAlaProCysThrLysIleLeu 222
 QY 139 CTGAGACTGGCGCGCTGAGAGCTCAACACCGCTGCTCCGAGAACCCAGAGATTCT 198
 DB 223 MetGlnPheCysAlaGlnIleGlnLeuTyrGlyValLeuAlaValArg----- 239
 QY 199 TCCCTTATACCCCTGGAGACATGCTCACCCCAAAAATAGCCTTACAGTCCGCTGGGC 258
 DB 240 -----ProValThrProSerLeuLeuValAspTyrPheMetGlyIleAlaGly 256
 QY 259 CTGGCTTACCTGACACAAAGAAACATCATCTTCTGTGACCTTAAGTCGACACATTCG 318
 DB 257 MetAsnTyrIleHisIleuHisLysIleIleHisArgAspLeuLysSerProAsnMetLeu 276
 QY 319 GTGTGCTCCCTTGACGCTCAAGAGACATCAACATCAACATCACTGCTGACGAGATTGCG 378
 DB 277 IleThrTyrAspAspVal-----ValLysIleSerAspPheGlyThrSer 291
 QY 379 AGGCAG---TCATTCCATGAGGCGCCCTAGGCGTGAGAGCACTGCTGCTACAGGCC 435
 DB 292 LysGlnLeuSerAspLysSerThrLysMetSerPheAlaGlyThrValAlaThrPheAla 311
 QY 436 CCAGAGATGAGGCGCTGATGTATGATGAGAGGTGATGATGATGATGATGATGATGATG 495
 DB 312 ProGlnValIleArgAsnGlnProValSerGlnLysValAspIleThrSerPheGlyVal 331
 QY 496 GTGCTTACAGATGCTGTGACAGAGAGCGCCT----- 528
 DB 332 ValLeuThrPheLeuLeuThrGlyGluIleProTyrLysAspValAspSerSerAlaIle 351
 QY 529 -----GCACTGGCGCCACCAAGCTCCAGATTCGCAAGAGCTGCCAAGGCAATCCG 582
 DB 352 IleThrPheLysValGlySerAsnSerLeuHisLeu----- 362
 QY 583 CCGGTCTGAGGAGCGCGAGAGAGTGCAGTTCGCGGCGAGCTGACGCGCTCATGATGAG 642
 DB 363 ProValProSerSerCysProAsp-----GlyPheLysIleLeuLeuArgGln 378
 QY 643 TCGTGGACATTAAGCCAGAGAGAGCGAGC-----CTGCGCGCTGCG 684
 DB 379 CysThrAsnSerLysProArgAsnArgProSerPheArgGlnIleLeuLeuHisLeuAsp 398
 QY 685 GTGGTGGCCAGATGAGAGAGAGCGAGCTTGGCCAGCTTATGATGATGATGATGATGATG 744
 DB 399 IleAlaSerAla-----AspValLeuSerThr 407
 QY 745 AAGCAGACAGCTTCTTCATCCAGGCGAG----- 777
 DB 408 ProGlnGlnThrTyrPheLysSerGlnAlaGlnTyrArgGlnGlnValLysLeuHisPhe 427
 QY 778 -----GASTACACCGCTGTGTT--- 795
 DB 428 GlnLysIleLysSerGlnGlnLysThrCysLeuHisArgLysGlnGlnIleuValMetArg 447
 QY 796 -----TGGGATCGAAGAAAGAGG 813
 DB 448 ArgArgGlyGlnPheArgHisIleAlaLeuAspIleArgGlnIstLysLysLeuGln 467
 QY 814 TCCAGAGAC-----TACAGCGGTGGAGAC-----ACAGAG 843
 DB 468 ArgAlaAsnAsnLeuTyrMetGlnLeuAsnAlaLeuMetLeuGlnLeuGlnThrGln 487
 QY 844 AAGGCTCATG-----GAGTGCAGAGAGATGCTGCGCTGGAGTGAAGTGAGC 894
 DB 488 ArgGlnLeuLeuArgArgAlaGlnAlaLeuGlnArgArgCysProGlyLeuLeuLysSer 507
 QY 895 TGGCAGCTCCAGAGTCCAGAGATCCCTGTGGAGAGAGCCAGGAGCCAGAAATTCATATC 954


```

OY 496 GTGCTACAGAGTCTGCTCAGACAGCGCCT----- 528
DB 332 ValLeuTPrgLLeuLeuThrlGylLeuLeuProTyrLysAspValAspSerSerAlaIle 351
OY 539 -----GCACTGGGCCACCACCAAGCTCCAGATTGCCAAGAACTGTCCAAAGGCATCCGC 582
DB 352 IleTPrgLylGylSerAsnSerLeuHisLeu----- 362
OY 583 CCGGTTCTGGGAGCCGAGAGAAAGTGCAGTCCGGCAGCTCAGCGCTCATGATGAG 642
DB 363 ProValProSerSerCysProAsp-----GlyPheLysIleLeuLeuArgGln 378
OY 643 TCGTGGACACTAAGCAGAGAAAGCAGCGC-----CTGGCCCTGTGC 684
DB 379 CysTPraAsnSerLysTPraArgAsnArgProSerPheArgGlnIleLeuLeuHisLeuAsp 398
OY 685 GTGTGAGCCAGATGAAGACCCGACTTTGCCACCTTCATGATGACTGTGTGGG 744
DB 399 IleAlaSerAla-----AspValLeuSerThr 407
OY 745 AAGCAGACAGCCTGTCTCATCCAGCGGCAG-----GAGTAC 783
DB 408 ProGlnGluThrTyrPheLysSerGlnAlaGluTPraGluGluValLysLeuHisPhe 427
OY 784 ACCGTGCTGTTTGGGATGA----- 804
DB 428 GluLysIleLysSerGluGluLysThrCysLeuHisArgLysGluGluLeuValMetArg 447
OY 805 ---AAGAGAGACTCCAGAAC----- 822
DB 448 ArgArgGluGluLeuArgHisAlaLeuAspIleArgGluHisTyrGluArgLysLeuGln 467
OY 823 -----TACACGGTGTGAACACA-----GAG 843
DB 468 ArgAlaAsnAsnLeuTyrMetGluLeuAsnAlaLeuMetLeuGlnLeuGluLysGln 487
OY 844 AAGGCGCTCATGAGAGTGCAGAGATGTC-----TGCCCTGGATGAGTGCAGC 894
DB 488 ArgGluLeuLeuArgArgGluGlnAlaLeuGluArgArgCysProLysLeuLeuLysSer 507
OY 895 TGCCAGCTCCAGTCCAGAGATCCCTGTGTGACAGCCAGGAGACCAAGAAATCTCATAC 954
DB 508 -HisProSerArgGlyLeuLeuHisGlyAsnThrMetGluLysLeuIleLysLysArg 527
OY 955 TACACCCCAAGGCGATGCGCCCTTAACACACCCCAAGGCGCTGTGATCTCCAGCT 1014
DB 527 nValProGlnLysLeuSer-----ProHisSerLysArgProAspLysLeuLysHis 544
OY 1015 GTGCTACAGCTGTTTGGCCGCTGCTGATTAAGAAAGAAATCTCACTGCTTACGC 1074
DB 544 rGluSer-----LeuLeuProLysLeuAspAlaAl 554
OY 1075 GGCCTGCGCGAGAGGCTGTGTGCTGTGTTCCCGTGGCGGAGCCCAAGAGACAGC 1134
DB 554 AleuSerGlyValGlyLeu----- 560
OY 1135 TGTCTACTGTGCTCAACACAGCCACAGCTCCCAAGTTCAGCATCCGCGATGAAGC 1194
DB 561 ---ProGlyCys-----ProLysGlyProProSerProGlyLysArgSerArg- 574
OY 1195 GCACGGCAGAACCCCTAACCCAGTGAAGGCGCATGAGGTGTGTAAACAGCGGCTGTGAGTC 1254
DB 575 -ArgGlyLysThrArgHisArg-----LysAlaSerAlaLys 586
OY 1255 TGTGTACAGAAATGGCGCGGCGCTCTGTGCATGAGTGTGCTCCCTGGAGATGTGACAG 1314
DB 586 scgLySerCysGlyAspLeuProGlyLeuArgAlaAlaLeuProGlnIleLysLysProGly 606
OY 1315 CGGCTGAGACCT-----ACATGGCC-----CTTCAGT 1344
DB 606 yLeuGlySerProGlyGlyLeuGlyValGlyProSerAlaIleTPraAspAlaCysProAl 626
OY 1345 GTTACGT-----CA 1353

```

```

DB 626 AleuArgGlyLeuHisHisAspLeuLeuLeuArgLysMetSerSerSerProAspLe 646
OY 1334 GTGCTGTGACGCTGTGAGGCGACAGGAGAGAGTCTGTGCTGCTGATGACA----- 1408
DB 646 uLeuSerAlaAlaLeuGlyAlaArgGlyArgGlyLysThrGlyAlaArgAspProG1 666
OY 1409 -AGCGCAACCTCTGTGTATGATGATGACATGACATGATGATGATGATGATGATGATG 1467
DB 666 ySerProProProProGlnGlyAspThrProProSerGluGlySerAlaProGlySerTh 686
OY 1468 TGCG-----GGTCCCGACCCCTCAGGAGACATGTTTCC 1503
DB 666 rSerProAspSerProGlyGlyAlaLysGlyGlyLysProProPro-----Pr 701
OY 1504 GTGCGCGCTTGTGACAGGAGAACCCCGGACGACGACGACGACGACGACGACGACG 1563
DB 701 oValGlyPro-GlyGlyGlyVal----- 708
OY 1564 GAGGAGACTCCATCGCGAGTGCATGATGATGATGATGATGATGATGATGATGATG 1623
DB 709 -GlyLeuGluGlyThrGlyArgGluGlyThrAlaGlyArgGlyLysAspArgAla 726
OY 1624 CTGATCCACAGGAATCACTACTACTACTACTACTACTACTACTACTACTACTACT 1683
DB 727 -----GlySerGlnHisLeuThrProAlaAlaLeuLeu-TyrArgAlaAlaVal 742
OY 1684 CCCCAGGCTGCGCAGGCTCCCTCAGGCTCCCGAGCTCCCGAGCAAGTCTTCCAGT 1743
DB 743 ThrArgSerGlnLysArgGlyLysSerSerGluGluGluGlyGlyValAspSerGlu 762
OY 1744 GTGCTTCTCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1803
DB 763 ValGluLeu-ProProSerGlnArg----- 770
OY 1804 GACAGTCTGACATGACCTGACCCCATGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1863
DB 771 -----TrpProGlnGlyProAsnMet----- 777
OY 1864 GCGGTGAGATCTCGCGCTCAGAGACCTCATTTGGTCCCGAGCGGCGTGAATGTT 1923
DB 778 -----ArgGlnSerLeuSerThrPheSerSerGluAsnProSerAspValGlu 793
OY 1924 ATCGTCATTTGCGTGGAGAGATTTCTGAAGCCCAAGCGGCGCGGATTCCTGTTA 1983
DB 794 -----GluGlyThrAlaSerGluProSerProSerG1 804
OY 1984 AAGCCCGAG 1993
DB 804 yThrProGlu 807

RESULT 9
A:39962
K:kinase-related transforming protein (abl) (EC 2.7.1.-) type I - mouse
N:Alternate names: protein-tyrosine kinase abl
C:Species: Mus musculus (house mouse)
C>Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 18-Jun-1999
C:Accession: A39962; A24773; S00771; A00626
R:Proc. Natl. Acad. Sci. U.S.A. 84, 8200-8204, 1987
A:Title: Nucleotide sequence of testis-derived c-abl cDNAs: Implications for testis-s
A:Reference number: A39962; MUID:88068561; PMID:3317402
A:Accession: A39962
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1123 <OBP>
A:Cross-references: GB:J02995; NID:G191566; PIDN:AA88241.1; PID:9309084
A:Experimental source: clone A16
R:Ben-Neriah, Y.; Bernard, A.; Paik, M.; Daley, G.O.; Baltimore, D.
Cell 44, 577-586, 1986
A:Title: Alternative 5' exons in c-abl mRNA.
A:Reference number: A90877; MUID:86133550; PMID:3512096
A:Accession: A24773

```

A: Molecule type: mRNA
 A: Residues: 1-118 <BER>
 R: Bernards, A.; Paskind, M.; Baltimore, D.
 Oncogene 2, 297-304, 1988
 A: Title: Four murine c-abl mRNAs arise by usage of two transcriptional promoters and a 5' noncoding region
 A: Reference number: S00771; MUID:88202920; PMID:3283651
 A: Accession: S00771
 A: Molecule type: DNA
 A: Residues: 1-26 <BER>
 A: Cross-references: EMBL:X07539; NID:949837; PIDN:CAA30411.1; PID:949838
 R: Wang, J.Y.; Ledley, F.; Goff, S.; Lee, R.; Groner, Y.; Baltimore, D.
 Cell 36, 349-356, 1984
 A: Title: The mouse c-abl locus: molecular cloning and characterization
 A: Reference number: A00626; MUID:84106840; PMID:6319018
 A: Accession: A00626
 A: Molecule type: DNA
 A: Residues: 85-127, 'C', 129-182 <MAN>
 A: Genetics:
 A: Gene: abl
 A: Superfamily: human protein-tyrosine kinase; protein kinase homology; SH2 homology;
 C: Keywords: alternative splicing; ATP; autophosphorylation; nucleus; phosphoprotein; pnc
 F: 68-116/Domain: SH3 homology <SH3>
 F: 127-217/Domain: SH2 homology <SH2>
 F: 240-500/Domain: protein kinase homology <KIN>
 F: 248-256/Region: protein kinase ATP-binding motif

Alignment Scores:
 Pred. No.: 1.03e-06 Length: 1123
 Score: 280.00 Matches: 186
 Percent Similarity: 34.69% Conservative: 95
 Best Local Similarity: 22.96% Mismatches: 285
 Query Match: 6.908 Indels: 244
 DB: 2 Gaps: 37

US-09-836-392-8_COPY_22_2205 (1-2184) x A39962 (1-1123)

QY 25 GATGCCAAGAACTTCCTCCGAGTCCGGCAGGAGCCAGATGTCAGCGCGTCGAG 84
 DB 276 AATPMTMGLN---ValGluPheLeuLysGluAlaValMetLysGluLeuLys 294
 QY 85 CACCCCTGACGTCGAGCGCTCATCGGACACATCCAC-----CCGCTCTGCTTGGC 138
 DB 295 HSPROASLeuValGlnLeuLeuGluValLysThrArgGluProPhePheTyrIle 314
 QY 139 CTGAGACTGGCGCGCTGACAGACCTCAACACCGCTGCTCGAG---AACGCCAGAGT 195
 DB 315 ThrGluPheMetThrTyrGluLysLeuLeuAspTyrLeuAlaGluLysAsnArgGln 334
 QY 196 TCTTCTTATACCCCTGGGACATGCTACCCAAATAAGCCCTACAGATCGGCTCG 255
 DB 335 ValSerAlaValValLeuLeuTyrMetAlaThr-----GlnIleSer 349
 QY 256 GGCCTGACCTGACGACAGAAACATCATCTTCTGACCTGAAGTCGACACAT 315
 DB 350 AlaMetGluTyrLeuLysLysAsnPheIleAspAspLeuAlaIaArgAsn 369
 QY 316 CTGGTGTGCTCCCTTACAGTCAAGAGCAGATCAATCAAGTATCTGACTCGGAT 375
 DB 370 LeuValGly-----GluAsnHisLeu---ValLysValAlaAspPheGlyLeu 384
 QY 376 TGGAGG-----CAGTCATTCAGATGAGCGGCC---CTAGCGGTGGAG 414
 DB 385 SerArgLeuMetThrGluAspThrTyrThrAlaHisAlaValAlaLysPheProIleLys 404
 QY 415 GGCACCTCTGCTGACAGAGCCGACAGATCAGGCTCGCATTCATATGATGAGAA 474
 DB 405 TrpThr-----AlaProGluSerLeuAlaLysAsnLysPheSerIleLysSer 420
 QY 475 GATATGTTCTCTATGAGATGTGCTCTACAGTCTGCTCA---GGACAGCGCCCTGCA 531
 DB 421 AspValTrpAlaPheGlyValLeuLeuTyrGluIleAlaThrTyrGlyMetSerProTyr 440
 QY 532 CTGGGACACACAGAGCTCCGATTCGCAAGAACTGTCCAAAGGCAATCCGCCGGTCTG 591

DB 441 ProGluIleAspLeuSerGlnValTyrGluLeuLeuGluLysAspTyrArg-----Met 458
 QY 592 GGGCAGCCGGAGAGAGTCAATGCTCCGGGACTGACGGGCTCATATGAGTGGGAC 651
 DB 459 GluArgProGluGlyCysPro---GluLysValTyrGluLeuMetArgAlaCysTrpGln 477
 QY 652 ACTAAGCCAGAGAACGACCGCTGCGCTGTCGGTG 687
 DB 478 TrpAsnProSerAspArgProSerPheAlaGluIleHisGlnAlaPheGluThrMetPhe 497
 QY 687 ----- 687
 DB 498 GlnGluSerIleSerAspLeuValGluLysGluLeuGlyLysArgGlyThrArgGly 517
 QY 688 -----GTGAGCCAGATGAAGACCCGACTTTCGCCACCTTATGTAAGTGGCTGT 741
 DB 518 GlyAlaGlySerMetLeuGlnAlaProGlu---LeuProThrLysThrArgThrCysArgArg 537
 QY 742 GGGAGCAGACAGCCTTCTCT-----CAT----- 766
 DB 537 GAlaAlaGluGlnLysAspAlaProAspThrProGluLeuLeuHisThrLysGlyLeuGln 557
 QY 766 ----- 766
 DB 557 YgluSerAspAlaLeuAspSerGluProAlaValSerProLeuLeuProArgLysGluArg 577
 QY 767 -----CCAGGCGCCAGAGATTC 783
 DB 577 gGlyProProAspGlySerLeuAsnGluAspLysArgLeuLeuProArgAspArgLys 597
 QY 784 ACCGTGCTTTTGGGATGAGAAAGAGAGCTCCAGAACTACAGGTGTGAACACAGAG 843
 DB 597 TAsnLeuPheSerAlaLeuIleLysLysLysLysMetAlaProThrProProLysArg 617
 QY 844 AAGGCTCATGAGAG-----TCAGAGATTCGCTGCCCTGG 882
 DB 617 gSerSerSerPheArgGluMetAspGlyGlnProAspArgGly----- 632
 QY 883 ATGAAGTGAGCTGCCAGCTCCAGATCCAGATCCCTGTGACACCCAGAGACAG 942
 DB 633 -----AlaSerGluAspSerArgGluLeuLysAsnGlyProProAlaLeuThr 649
 QY 943 AAATCTACATCTACACCTCAAGGCAATGCGCCCTTAACACCCACAGCGCTTG 1002
 DB 649 TserAspAlaAlaGluProThrLysSer----- 658
 QY 1003 GATATCCAGCTGTGCTCACTCTTCTTGGCCCTGCTATTAAAGAATTCTTAC 1062
 DB 659 -----ProLysAlaSerAsnGlyAlaGlyValProAsnGlyAlaPheArgGluProGlyAs 677
 QY 1063 CTGGCTTACGGGCGCTCCCGATGGGCTTGCGCTGTTCCCTGCGGGGCGAC 1122
 DB 677 nserGlyPheArgSerProHisMet-----TrpLysLysSer 690
 QY 1123 CCAAGAGACAGTGCCTACCTGCTGCTCAGACACAGACAGGTCAGTCAAGTCA 1182
 DB 690 ThrLeuThrGlySerArgLeuAlaAlaAlaGluGlnGluSerGlyMetSerSerIle 710
 QY 1183 GCGGATG-----AAGAGCAGCGGAGAACCCCTACCCAGTG 1218
 DB 710 SarGlyPheLeuArgSerCysSerAlaSerCysMetProHisGlyAlaArgAspThrGlu 729
 QY 1219 AAGCCATGAGAGTGTCAACAGCGGCTCTGAGGTGTGACAGCATAGCGGGGCTTC 1278
 DB 730 -----TrpArg---SerValThrLeuProArgAspLeuProSerAlaGlyLysInp 746
 QY 1279 CTTCGATGAGCTGCTCCCTCCCTGAGATCTGCAGGCGCTGAGCCCTTACAGGCC 1336
 DB 746 eAspSerSerThrPheGlyGlyHisLysSerGlu-----LysProAlaLeuProArg 763
 QY 1337 -----CC 1338

```

Db 763 glysargThrSerGluSerArgSerGluGlnValAlaLysSerThrAlaMetProLeuPr 783
QY 1339 TCCATGCTTACTCAGTCTGCTGACCTGTGAGGCGAGGAGGAGGAGCTGCTGCTGC 1398
Db 783 ocllythrLeuLysLysAsn-----GluGlnAlaAlaGluGlnLys 797
QY 1399 CTGATGACAAAGCAGCCTGTGATGTACACTCCACCTCCACCTCCAGCTGTGCTGC 1458
Db 797 elysasphThrGluSerSerPro-GlySerSerPro-----ProSerLeuThrP 813
QY 1459 CGGTACTTCTGCGG-----GATCCCGAGCCCTCCAGGAGGAGCTGTTTCC 1503
Db 813 rolyLeuLeuArgArgGlnValThrAlaSerProSerSerGlyLeuSerHisLysGluG 833
QY 1504 GTCGC-----GCCCTTGACAGGAAACCCCGGACGACGACGACGACGACGAC 1551
Db 833 lualathrLysLysSerAlaSerGlyMetLysThr-ProAlaThrAlaGluProAlaPro 852
QY 1552 CCAAAGTCTCTGAGGAGGAGCTCCATCCGAGCTGAGCATCATGACATGAGAGAGT 1611
Db 853 ProSerAsnLysValGlyLeuSerLysAlaSer-----SerGluGlnMet 867
QY 1612 GGCAGCAGATCCTGTATCCACAGAGATCTACTACTGCTCATGCTCTCTAC 1671
Db 868 -----ArgValArgArgHisLysHisSer----- 875
QY 1672 TCCATATCCCGACCCGCGAGCTGCGCAG----- 1701
Db 876 SerGluSerProGlyArgAspLysGlyArgLeuAlaLysLeuLysProAlaProPro 895
QY 1702 -----TCCCTTCAGCCTCCCGCAGCTCCCGCAGCAGT 1734
Db 896 ProProAlaCysThrGlyLysAlaGlyLysProAlaGlnSerProSerGlnGluAlaGly 915
QY 1735 TCTTCAGTGTGCTTCTTCCACCCGCTGCGAGAC-----TCAGACATGCTACTAC 1788
Db 916 GluAlaGlyLysProThrLysThrLysCysThrSerLeuAlaMetAspAlaValAlaThr 935
QY 1789 -----CCGCTGCTGCTCCGACAGGCTGAGCATGAGCTGAGCCCGCAGGAC 1836
Db 936 AspProThrLysAlaGlyPro-ProGlyLysLysLeuAlaGlyLysProValProSerVal 955
QY 1837 GGGAGACCTTACAGCAGCAGCTGACGAGCGGTGAGATCTCCGCTCAGACAGCTCAT 1896
Db 955 lProLysProGln-----SerThrAlaLysProProGlyThrProThrSerProValSer-- 973
QY 1897 TGGGTCCCGAGGCGGTGAGATGTATCGTCATTCGCTGAGAGAGATTCGAGCC 1956
Db 974 -----ThrProSerThrAlaProAlaProSerProLeuAlaGlyAspGln-----GlnPr 990
QY 1957 CAGCGGCGCGAGTATTCGCG 1978
Db 990 oSerSerAlaAlaPheIlePro 997

RESULT 10
J00229
mixed-lineage protein kinase 1 - human
C.Species: Homo sapiens (man)
C.Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C.Accession: S32467; J00229
R.Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A.Title: Identification of a new family of human epithelial protein kinases containing
A.Reference number: S32467; M01D:93238756; PMID:8477742
A.Accession: S32467
A.Molecule type: mRNA
A.Residues: 1-394 <DO>
C.Genetics:
A:Gene: GDB:MLK1
A:Cross-references: GDB:141921; OMIM:600136
A:Map position: 14q24.3-14q31
A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k

```

```

F.1-269/Domain: protein kinase homology <KIN>
F.1-268/Domain: catalytic <CAT>
F.9-17/Region: protein kinase ATP-binding motif
F.289-310/Region: leucine zipper motif
F.324-345/Region: leucine zipper motif
F.354-368/Region: basic

Alignment Scores:
Pred. No.: 1,18e-06 Length: 394
Score: 279.50 Matches: 87
Percent Similarity: 46.05% Conservative: 47
Best Local Similarity: 29.90% Mismatches: 108
Query Match: 6.89% Indels: 50
DB: 2 gaps: 9

US-09-836-392-8_COPY_22_2205 (1-2184) x J00229 (1-394)
QY 7 AGCAGCTTCCGCGGACCCAGCTCCATGCAAGTCTCCGAGCTCCGCGAGGCGCAGC 66
Db 33 ArgHisAspProAspGlnAspLysPheSerGlnThrIleGluAsnValArgGlnLys 52
QY 67 ATGTGACGCGCTGACGAGCAGCTGATGCTGCGCTCATCGCATCCAGTCCACCG 126
Db 53 LeuPheAlaMetLeuLysHisProAsnIleIleAlaLeuArgLysValCysLeuLysGlu 72
QY 127 -----CTCTGCTTGGCCCTGAGCTCGCGCGCTCAGCAGGCTCAACACGCTGCTCC 180
Db 73 ProAsnLeuCysLeuValMetGluPheAlaArgGlyGlyProLeuAsnArgValLeuSer 92
QY 181 GAGAACGCGCAGAGATCTCTTATACCCCTGAGCAGACATCTGACCCCAAAATA-- 237
Db 93 -----GlyLysArgIleProProAspLysLeu 101
QY 238 -----GCCTACAGATGCTGCTGCGCTGCGCTTACCTGACACAGAA-- 279
Db 102 ValAsnThrPalaValGlnIleAlaArgGlyMetAsnTyrLeuHisAspLysAlaLeuVal 121
QY 280 AACATCATCTTCTGAGCTGAGCTGAGCGGACACATCTTCTGCTGCTGCTGCTGCT 330
Db 122 ProIleIleHisArgAspLeuLysSerSerAsnIleLeuIleLeuGlnLysValLysAsn 141
QY 331 ---GACGTCAGAGCAGACATCAACATCAAGTATCTGATGAGGATTCAGGAGTCA 387
Db 142 GlyAspLeuSerAsnLysLeu-----LeuLysIleThrAspPheGlyLeuAlaArgLysThr 160
QY 388 TTCATGAGGCGCGCTGAGCGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
Db 161 HisArgThrThrLysMetSerAlaAlaGlyThrTyrAlaThrMetAlaProGluValIle 180
QY 448 CCTCGATGTATGATGATGAGAGAGTATGATGCTCTCTATGAGATGCTGCTACGAG 507
Db 181 ArgAlaSerMetPheSerLysLysSerAspValTyrPheThrGlyValLeuLeuThrPglu 200
QY 508 TTGCTGTACAGACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558
Db 201 LeuLeuThrGlyGluValProPheArgGlyIleAspGlyLeuArgValAlaLysThrGlyVal 220
QY 559 -----AAGACCTGTCAGAGGCGATCCGCGCTTGGGCGAGCGGAGAGAGTGCAG 612
Db 221 AlaMetAsnLysLeuAlaLeuProIleProSerThrCysProGluPro----- 236
QY 613 TTCGCGGAGCTCCAGCGCTGATGATGAGTCTGCTGAGACACTRACCGAGAGAGAGCG 672
Db 237 -----PheAlaLysLeuMetGluAspCysThrAspProAspProHisSerArg-- 252
QY 673 CTGCGCTTCTGCTGCTGAGCAGATGAGAGCGAGCTTTCACACCTTCATGATGAA 732
Db 253 -----ProSerPheThrAsnIleLeuAspGln 261
QY 733 CTGTGCTGTGGAGACAGACCTTCTTCTC-ATCCAGGCGCAGAGTACACCGTGT 791
Db 262 LeuThrThrIleGluGlnSerLysPhePheGlnMetProLysAspSerPheHisCysLeu 281

```

QY 792 GPTTGGATGAGAAAGAGAGCCAGCACTA 824
 DB 282 GlnAspAsnTrpLysHisGluLeuGlnMet 292
 RESULT 11
 D84555
 Probable protein kinase (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: D84555
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 Eues, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: D84555
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-546 <STO>
 A:Cross-references: GB:AE002093; NID:96598802; PIDN:AAB80785.2; GSPDB:GND0139
 A:Gene: At2g17700
 A:Map position: 2
 Alignment Scores:
 Pred. No.: 1,29e-06 Length: 546
 Score: 278.50 Matches: 71
 Percent Similarity: 52.02% Conservative: 45
 Best Local Similarity: 31.84% Mismatches: 90
 Query Match: 6.86% Indels: 17
 DB: 2 Gaps: 8
 US-09-836-392-8_COPY_22_2205 (1-2184) x D84555 (1-546)
 QY 46 GAGTTCGGCAGAGAGCCAGCTGCAAGCGCTGACAGACCCCTGCATCGTGGCGCTC 105
 DB 327 GlnPheSerGlnGluValPheIleMetArgLysValAlaHisLysAsnValValGlnPhe 346
 QY 106 ATCGGCATCAGCATCCAC-----CCGCTCTGCTTGGCCCTGAGAGTCCGCCGCCCTCAGC 159
 DB 347 LeuGlyAlaCysThrArgSerProThrLeuGlyIleValThrGlnPheMetAlaArgGly 366
 QY 160 AGCGCTGAACCGCTGCTGTCGAGAGAGCCAGAGATTCTCTTATACCCCTGGGACAC 219
 DB 367 SerIleLysAspPheLeuHis-----LysGlnLysCysAlaPheLysLeuGlnThr 383
 QY 220 ATGCTGACCCAAATAATAGCCTACAGATGCGCTGCGGCTGCTACCTGACCAAGAA 279
 DB 384 LeuLeu-----LysValAlaLeuAspValAlaLysGlyMetSerLysLeuHisGlnAsn 401
 QY 280 ACATCATCTCTCTGTGACCTGGAAGTGGGACACATCTGTGTGGTCCCTTGACGTCAAG 339
 DB 402 AsnIleLeuHisArgLysPheLysLysThrAlaAsnLeuLeu-----MetAsp 416
 QY 340 GAGCAGATCAACATCAAGTATCTGACTAGCGGATTTGAGAGCAGATCATTCATGAGGCG 399
 DB 417 GlnHisGlyLeuValLysValAlaAspPheGlyValAlaArgValGlnIleGluSerGly 436
 QY 400 GCCCTAGCGGTGAG---GGCACTCTGCTGCTACAGGCCCGCCAGAGATCAGGCTCGCAT 456
 DB 437 ValMetThrAlaGluThrGlnGlyThrArgTrpMetAlaProGluValIleGlnHisLys 456
 QY 457 GWTATGATGAGAAAGGTAGTATGTTCTCTATAGGAATGCTGCTACGATGCTGCTCA 516
 DB 457 ProLysHisLysIleValAlaAspPheSerLysValAlaIleValLeuTrpGluLeuLeuThr 176
 QY 517 GAGACAGGCGCTGACGCTGGGCGCAACACAGCTCCAGATGCGC---AAGAGCGTGTCCAG 573
 DB 477 GlyAspIleProLysAlaPheLeuThrProLeuGlnAlaValAlaGlyValGlnLys 496
 QY 574 GGCATCCGCCCGGTTCTGGGGGACCGCGAGAAAGTGCAGTCCGCGCAGCTGCAGCGCTC 633
 DB 574 GGCATCCGCCCGGTTCTGGGGGACCGCGAGAAAGTGCAGTCCGCGCAGCTGCAGCGCTC 633

DB 497 GlyLeuArgProLysIle-----ProLysLysThrHis---ProLysValLysGlyLeu 513
 QY 634 ATGATGCACTGCTGGGACACTAGACGAGAGAGACCGGCTGCGCTGCGGTGAGC 693
 DB 514 LeuGlnArgCysTrpHisGlnAspProGluGlnArgProLeuPheGluGlnIleGlu 533
 QY 694 CAGATGAAG 702
 DB 534 MetLeuGln 536
 RESULT 12
 FOMVGM
 gag-abl polyprotein - Abelson murine leukemia virus
 N:Contains: amino end of core shell protein p30; core protein p15; inner coat protein
 C:Species: Abelson murine leukemia virus
 A:Note: host kus sp. (mouse)
 C:Date: 14-Nov-1983 #sequence_revision 09-Sep-1994 #text_change 11-Jun-1999
 C:Accession: A03931; A00627; A93955
 R:Reddy, E.P.; Smith, M.J.; Srinivasan, A.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3623-3627, 1983
 A:Title: Nucleotide sequence of Abelson murine leukemia virus genome: structural siml
 A:Reference number: A93955; MUID:83221648; PMID:6304726
 A:Accession: A03931
 A:Molecule type: DNA
 A:Residues: 1-981 <RED>
 A:Cross-references: GB:J02009; NID:9331887; PIDN:AAA46471.1; PID:9331888
 A:Note: the authors translated the codon GGA for residue 186 as Glu
 R:Reddy, E.P.; Smith, M.J.; Srinivasan, A.
 Proc. Natl. Acad. Sci. U.S.A. 80, 7572, 1983
 A:Reference number: A93980
 A:Contents: annotation; erratum; residues 588-746
 C:Genetics:
 A:Gene: gag-abl
 C:Superfamily: Abelson murine leukemia virus gag-abl polyprotein; protein kinase homo
 C:Keywords: ATP; core protein; oncogene; phosphotransferase; polyprotein; transformin
 F:1-131/Product: core protein p15 #status predicted <p15>
 F:132-215/Product: inner coat protein p12 #status predicted <p12>
 F:216-235/Region: amino end of core shell protein p30
 F:248-338/Domain: SH2 homology <SH2>
 F:361-621/Domain: protein kinase homology <KIN>
 F:369-377/Region: protein kinase ATP-binding motif
 F:392/Active site: Lys #status predicted
 Alignment Scores:
 Pred. No.: 2.17e-06 Length: 981
 Score: 273.50 Matches: 156
 Percent Similarity: 34.83% Conservative: 68
 Best Local Similarity: 24.15% Mismatches: 226
 Query Match: 6.74% Indels: 195
 DB: 1 Gaps: 30
 US-09-836-392-8_COPY_22_2205 (1-2184) x FOMVGM (1-981)
 QY 25 GATGCAATGAGAACTTCTCCGAGTCCGGAAGGAGCCAGATGCTGACAGCGCTGCAG 84
 DB 397 AspThrMetGlu---ValGlnGluPheLeuLysGlnAlaValMetLysGluIleLys 415
 QY 85 CAGCCCTGATGCTGGCGCTCATGCGCATGCGCATGCGATCCAC-----CCGCTGCTGCGC 138
 DB 416 HisProAsnMetValAlaGlnLeuLeuGlyValCysThrArgGluProProPheTrpIleIle 435
 QY 139 CTGAGACTCGCGCGGCTGCAAGCCTCAACACCGCTGTCGAG -AAGCGTACAGAT 195
 DB 436 ThrGlnPheMetThrTrpGlnAsnLeuLeuAspTrpLeuArgGluCysAsnAlaGlnGlu 455
 QY 196 TCTCTCTTATACCCCTGCGGACATGCTCCACCCAAATAATAGCTTACAGATGCGCTCG 255
 DB 456 ValSerAlaValAlaLeuLeuTrpMetAlaThr-----GlnIleSerSer 470
 QY 256 GCGCTGCGCTACCTGCAAGAAACATCATCTCTGTGACCTGAAGTGGACACAT 315
 DB 471 AlaMetLysTrpLeuGlnLysLysAsnPheIleHisArgAspLeuAlaAlaArgAsnCys 490


```

Db 336 ILeGyLaAcSthArgSerProAsnLeuCyStLeValThrGluPheMetThrArgGly 355
QY 160 AGCCCAACACCCGCTCTCCGAGAACCCAGATTCCTCTTATATCCCTCGGACAC 219
Db 356 SerIleThrAspPheLeuHisLysHis-----LysGlyValPheLysIleInsSer 372
QY 220 ATGCTCACCCAAATAATAGCCATACAGATCCGCTCGGCTCTGCTTACTGACACAGAA 279
Db 373 LeuLeu-----LysValAlaLeuAspValSerLysGlyMetAspTyrLeuHisGlnAsn 390
QY 280 AACATCATCTTCTGCTGAGTGCAGACAAACATTCCTGCTGCTGCTGCTGCTGCTGCTG 339
Db 391 AsnIleIleHisArgAspLeuLysThrAlaAsnLeu-----MetAsp 405
QY 340 GAGCAGCATCAACATCAACATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 399
Db 406 GluHisGluValAlaValAlaAspPheGlyValAlaArgValGlnThrGluSerGly 425
QY 400 GCCCTAGGCGTGAG-----GGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 456
Db 426 ValMetThrAlaGluThrGlyThrGlyArgTyrMetAlaProGluValIleGluHisLys 445
QY 457 GATATGATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 516
Db 466 GlyGluLeuProTyrSerTyrLeuThrProLeuGlnAlaValGlyValGlnLys 485
QY 574 GGCATCCCGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 633
Db 486 GlyLeuAspProLysIle-----ProLysGluThrHis--ProLysLeuThrGluLeu 502
QY 634 ATGATGAGATGCTGAGACATGACATGACATGACATGACATGACATGACATGACATGAC 693
Db 503 LeuGluLysCysTyrProGlnAspProAlaLeuArgProAsnPheAlaGluIleGlu 522
QY 694 CAGATGAAGAGCCGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 753
Db 523 MetLeuAsnGlnLeuIleArgGluValAlaAspLeuSerLeuHisLysAspHisGly 542
QY 754 GCTCTCTTCTCA 765
Db 543 GlyTyrPheSer 546

RESULT 14
T12956
hypothetical protein T6H20.50 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 22-Oct-1999
C:Accession: T12956
R:Choi, N.; Robert, C.; Brotlier, P.; Wincker, P.; Cattivello, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, July 1999
A:Reference number: Z17586
A:Accession: T12956
A:Molecule type: DNA
A:Residues: 1-1171 <CHO>
A:Cross-references: EMBL:AL096859; GSPDB:GN00061; ATSP:T6H20.50
A:Experimental source: cultivar Columbia; BAC clone T6H20
C:Genetics:
A:Gene: ATSP:T6H20.50
A:Map position: 3
A:Introns: 780/1; 857/3; 881/3; 932/3; 987/2; 1036/3; 1077/3; 1107/1

Alignment Scores:
Pred. No.: 3,01e-06 Length: 1171
Score: 270.50 Matches: 80
Percent Similarity: 49.80% Conservative: 42
Best Local Similarity: 32.65% Mismatches: 100
Query Match: 6.67% Indels: 23
DB: 2 Gaps: 9

```

```

US-09-836-392-8_COPY_22_2205 (1-2184) x T12956 (1-1171)
QY 16 CGGGCCACCATGCCAGAGAACTTCTCCAGATTCGCGGACGAGCCAGCATGCTGAC 75
Db 925 LysProSerGluGlnGluArgMetLeuAspPheThrAsnGluAlaGlnAsnLeuAla 944
QY 76 GCGCTGACGACCCCTGATGCGGCTCATCGGATGACATGACATGACATGACATGACATG 123
Db 945 GlyLeuHisLysProAsnValAlaAlaPheTyrGlyValAlaAspSerProGlyGly 964
QY 124 CCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
Db 965 SerValAlaThrValThrGluTyrMetValAsnGlySerLeuArgAsnAlaLeuGlnLys 984
QY 184 AACGCCAGAGATTTCTCTTATACCCCTGGGACATGCTCACCCAAATAATAGCTTAC 243
Db 985 AsnValArgAsn-----PheAspArgCysLysArgGlnLeuIleAlaMet 999
QY 244 CAGATCGCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303
Db 1000 AspIleAlaPheGlyMetGluTyrLeuHisGlyLysLysIleValHisPheAspLeuLys 1019
QY 304 TCGGACAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 1020 SerAspAsnLeuLeuValAsnLeuArgAspProHisArgProIle---CysLysValGly 1038
QY 364 GACTACGGGATTTCTGAG-----CAGTCATTCATGAGGCGGCTGAGCTGAG 414
Db 1039 AspLeuGluLysSerLysValLysCysGlnThrLeuIleSerGly-----GlyValArg 1056
QY 415 GGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 468
Db 1057 GlyThrLeuProTyrMetAlaProGluLeuLeuAsnGlyThrSerSerLeuValSerGlu 1076
QY 469 AAGTAGATATGTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 528
Db 1077 LysValAspValPheSerPheGlyLysValLeuThrGluLeuPheThrGlyGluGluPro 1096
QY 529 ---GCATGCGCCACACACCATGCTCCAGATTCGCGGACGAGCTGAGGCTGAGGCTG 585
Db 1097 TyrAlaAspLeuHisTyrGlyValAlaIleLeuGlyGlyIleValSerAsnThrLeuArgPro 1116
QY 586 GTTCTGGGCGACCGGAGAGTGCATGCGGCGAGCTGCGGAGCTGAGGCTGAGGCTGAG 645
Db 1117 GlnIleProAspPheCysAspMetAspTyrLys-----LeuLeuMetIuArgCys 1133
QY 646 TGGGACATGACCCAGAGAGCGACCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
Db 1134 TrpSerAlaGluProSerGluArgProSerPheThrGluIleValAlaGlnIleuArg--- 1152
QY 706 CCGACTTTGCGCAC 720
Db 1153 ---ThrMetAlaThr 1156

RESULT 15
TVHUA
protein-tyrosine kinase (EC 2.7.1.112) abl - human
N:Alternate names: kinase-related transforming protein p150
C:Species: Homo sapiens (man)
C:Date: 03-Aug-1984 #sequence_revision 17-Nov-1995 #text_change 04-Feb-2000
C:Accession: S08519; A25582; A00625; E38268
R:Fainstein, E.; Elnat, M.; Gokkel, E.; Marcelle, C.; Croce, C.M.; Gale, R.P.; Canaan
Oncogene 4, 1477-1481, 1989
A:Title: Nucleotide sequence analysis of human abl and bcr-abl cDNAs.
A:Reference number: S08519; MUID:90082420; PMID:2687768
A:Accession: S08519
A:Molecule type: mRNA
A:Residues: 1-1130 <FA1>
A:Cross-references: EMBL:X16416
R:Shitvelman, E.; Lifshitz, B.; Gale, R.P.; Roe, B.A.; Canaan, E.
Cell 47, 277-284, 1986
A:Title: Alternative splicing of RNAs transcribed from the human abl gene and from th

```

A:Reference number: A25582; MUID:87028219; PMID:3021337
 A:Accession: A25582
 A:Molecule type: mRNA
 A:Residues: 1-139, 'P', 141-158, 'S', 160-444, 'R', 446-458, 'K', 460-718, 'V', 720-836, 'W', 838-86
 'S', 1104-1130 <SH2>
 A:Note: the authors translated the codon GAG for residues 279, 282, 527, 549, and 581 as
 Argofen, J.; Helstetkamp, N.; Reynolds Jr., F.H.; Stephenson, J.R.
 Nature 304, 167-169, 1983
 A:Title: Homology between phosphotyrosine acceptor site of human c-abl and viral oncogen
 A:Reference number: A00625; MUID:83245023; PMID:6191223
 A:Accession: A00625
 A:Molecule type: DNA
 A:Residues: 360-423, 'GK', 426 <GRO>
 R:Partners: J.; Mekelae, T.P.; Altalo, R.; Lehaesialho, H.; Altalo, K.
 Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990
 A:Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.
 A:Reference number: A38268; MUID:91062389; PMID:2247464
 A:Accession: E38268
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 365-420 <PAR>
 C:Genetics:
 A:Gene: GDB:ABLI
 A:Cross-references: GDB:119640; OMIM:189980
 A:Map position: 9q34.1-9q34.1
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: human protein-tyrosine kinase ab1; protein kinase homology; SH2 homology;
 C:Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferase;
 F:68-116/Domain: SH3 homology <SH3>
 F:127-217/Domain: SH2 homology <SH2>
 F:240-500/Domain: protein kinase homology <KIN>
 F:248-256/Region: protein kinase ATP-binding motif
 F:146,770/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:127/Active site: Lys status predicted
 F:393/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Alignment Scores:
 Pred. No.: 3 19e-06 Length: 1130
 Score: 270.00 Matches: 198
 Percent Similarity: 36.05% Conservative: 94
 Best Local Similarity: 24.44% Mismatches: 283
 Query Match: 6.65% Indels: 236
 DB: 1 Gaps: 46

US-09-836-392-8_COPY_22_2205 (1-2184) x TVHUA (1-1130)

```

QY 25 GATGCCATGAGACACTTCTCCGAGTCCGCGAGAGCCGATCTGCACCGCTGCAG 84
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 276 AspThrMetGlu---ValGluGluIupheLeuLysGluAlaAlaValMetLysGluLeu 294
QY 85 CACCCCTGCATCGTGGCGCTCATCGGCATCATCCAC-----CCGCTCTGCTTCCGC 138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 295 HisProanLeuValGluLeuLeuGluValLysThrArgGluProProPheTyrIleLe 314
QY 139 CTGGAGTCGGCGCGCTCAGACGCTCAACACCGCTGCTCCGAG---AACCGACAGAT 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 315 ThrGluPheMetThrTyrGluYasnLeuLeuAspTyrLeuArgGluLysAsnArgGlu 334
QY 196 TCTTCCTTATACCCCTGGACACATGCTCACCACCAAAATAGCTTACAGATCGCTCG 255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 335 ValAsnAlaValValLeuLeuTyrMetAlaThr-----GlnIleSerSer 349
QY 256 GGCCTGGCTCATCTGCACAGAAAACATCATCTTCTGACCTGAGTGGACACAT 315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 350 AlameTgIuTyrLeuGluLysAsnPheIleHisArgAspLeuAlaIleArgAsn 369
QY 316 CTGGGTGTGCTCCTGCATGACATCAAGACATCAACATTAACATCTACTACAGGAT 375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 370 LeuValIglY-----GluAsnHisLeu---ValIleValAlaIleAspPheGlu 384
QY 376 TCGAGG-----CAGTCATTCATGAGCGGCC---CTAGCGCGTGAG 414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 385 SerArgLeuMetThrGluAspThrTyrThrAlaHisAlaGluAlaLysPheProIleLys 404
QY 415 GGCACCTCTGCTACAGAGCCCGACAGATCAAGCCCTCCATTTGATATGAGAGTA 474
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 405 TrpThr-----AlaProGluSerLeuAlaTyrAsnLysPheSerIleLysSer 420
QY 475 GATATGTTCTCTATGAGAAATGCTGCTTACAGATTCGTCTCA---GGACAGCGCCCTCA 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 AspValIlePheAlaPheGluValLeuLeuTyrGluIleAlaThrTyrGluMetSerProTyr 440
QY 532 CTGGCCACACACACCTCAGATTCGCAAGAGAGCTGTCAGAGGATCCCGGCTTCTG 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 441 ProGluIleAspLeuSerGlnValTyrGluLeuLeuGluLysAspTyrArg-----Met 458
QY 552 GGCACCGCGAGAGAGATCCAGTCCGCGACTGCAGCGCTCATGATGAGTGTCTGGAC 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 459 GluArgProGluGluCysPro---GluLysValTyrGluLeuMetArgAlaCysTrpIle 477
QY 652 ACTAGCCAGAGAACAGACCGCTGCGCTGCGGTGGAGGAGCCAGATGAGACCCGACT 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 478 TrpAsnProSerAspArgPro---SerPheAlaGluIleHisGln-----Ala 492
QY 712 TTTCGACCTTCATGATGATGACTGTCTGTGGAGACAGACCCCTTCTCATCCAG 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 493 PheGluThrMetPheGluGlu---SerSerIleSerAspGluValGluLysGlu-LeuGlyL 512
QY 772 GGCAGAGATCACCGCTGTGTTGGATGGAAGAGAGACTCCAGACTA-----824
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 512 YsgIingIyVal---ArgGluAlaValSerThrLeuLeuGluAlaProGluLeuProThrL 531
QY 825 -----CACGCTGTGAACACAGA-----GAGGCG 849
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 531 YsthrArgThrSerArgAlaAlaGluHisArgAspThrThrAspValProGluMetP 551
QY 850 CTCATGGA-----GATCGACAGATGTGCTCCCTGGGATGAGAGTGCAGC 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 551 RohIleSerLysGluGluGluGluSerAspProLeuAspHisGlu-----ProA 567
QY 901 CTCACAGTCCAGAGATCCCTGTGGACAGCCAGGACAGACAAATCTACTATCACAC 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 567 IalValSerProLeuLeuPro---ArgLysGluArgGlyPro-----P 580
QY 961 CTCAGGGCATGTGCCCTTAACACACCCAGAGCGCTGATACACAGCTGTGTC 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 580 Roglu-Gly-----GluLeuAsnGluAspGluArgLeuLeuProLysAspLysThr 597
QY 1021 ACCTGCTTCTGGCCGTG-----CCTGTTATTAA 1050
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 598 AsnLeuPheSerAlaLeuIleLysLysLysLysLysThrAlaProThrProProLysArg 617
QY 1051 AAGAAATCTCCTACCTGCTTACCGGCG-----1077
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 618 SerSerSerPheArgGluMetAspGluProGluArgArgGlyAlaGluGlu 637
QY 1078 -----CTGCCGATGAGCGCTTGTGCTGTGTTCCCGTGGCGGACCCCAAG 1128
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 638 GlyArgAspIleSerAsnGluValAlaLeuAlaPhe-----ThrProLeu 651
QY 1129 GACAGCTGCTCTACCTGTGCTCACACACAGCCACAGT-----CCAAGT 1176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 652 AspThr-AlaAspProAlaLysSerProLysProSerAsnGluAlaGluValProAsnGlu 671
QY 1177 AGCATCGCGGATGAGACGCGCGC-----AGAACCCCTACCCAGTGAAGCGCATG 1227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 671 ValAlaLeuArg---GluSerGluGluSerGlyPheArgSerPro-----HisLeuTr 687
QY 1228 GAGTGTGCACACAGCGGCTGTGAGTGTGTACACATGGC-----1270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 687 PylsLysSerSerThrLeuThrSerSerArgLeuAlaThrGluGluGluGlyGly 707
QY 1271 -----CGGCTCTCTTCATCTGAGTGTGCT-----1297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ySerSerSerLysArgPheLeuArgSerCysSerAlaSerCysValProHisGlyAlaLys 727

```